

UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

ESTRATÉGIA PARA PREDIÇÃO DE CONSUMO DE ENERGIA ELÉTRICA DE CURTO  
PRAZO: UMA ABORDAGEM BASEADA EM DENSIFICAÇÃO COM *MEAN SHIFT*  
PARA TRATAMENTO DE DIAS ESPECIAIS

LIVIANE PONTE RÊGO

TD 18/2016

UFPA/ITEC/PPGEE  
Campus Universitário do Guamá  
Belém – Pará – Brasil  
2016

UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

LIVIANE PONTE RÊGO

ESTRATÉGIA PARA PREDIÇÃO DE CONSUMO DE ENERGIA ELÉTRICA DE CURTO  
PRAZO: UMA ABORDAGEM BASEADA EM DENSIFICAÇÃO COM *MEAN SHIFT*  
PARA TRATAMENTO DE DIAS ESPECIAIS

TD 18/2016

UFPA/ITEC/PPGEE  
Campus Universitário do Guamá  
Belém – Pará – Brasil  
2016

UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

LIVIANE PONTE RÊGO

ESTRATÉGIA PARA PREDIÇÃO DE CONSUMO DE ENERGIA ELÉTRICA DE CURTO  
PRAZO: UMA ABORDAGEM BASEADA EM DENSIFICAÇÃO COM *MEAN SHIFT*  
PARA TRATAMENTO DE DIAS ESPECIAIS

Tese submetida à Banca Examinadora do  
Programa de Pós-graduação em Engenharia  
Elétrica da UFPA para obtenção do Grau de  
Doutor em Engenharia Elétrica na área de  
Computação Aplicada.

UFPA/ITEC/PPGEE  
Campus Universitário do Guamá  
Belém – Pará – Brasil  
2016

## FICHA CATALOGRÁFICA

UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

ESTRATÉGIA PARA PREDIÇÃO DE CONSUMO DE ENERGIA ELÉTRICA DE CURTO  
PRAZO: UMA ABORDAGEM BASEADA EM DENSIFICAÇÃO COM *MEAN SHIFT*  
PARA TRATAMENTO DE DIAS ESPECIAIS

LIVIANE PONTE RÊGO

TESE DE DOUTORADO SUBMETIDA À AVALIAÇÃO DA BANCA EXAMINADORA  
APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM  
ENGENHARIA ELÉTRICA DA UNIVERSIDADE FEDERAL DO PARÁ E JULGADA  
ADEQUADA PARA OBTENÇÃO DO GRAU DE DOUTOR EM ENGENHARIA  
ELÉTRICA NA ÁREA DE COMPUTAÇÃO APLICADA.

APROVADA EM / / 2016

BANCA EXAMINADORA:

---

Prof. Dr. Ádamo Lima de Santana (ORIENTADOR – PPGEE/UFPA)

---

Prof. Dr. Carlos Renato Lisboa Francês (CO-ORIENTADOR – PPGEE/UFPA)

---

Prof. Dr. Diego Lisboa Cardoso (MEMBRO PPGEE – UFPA)

---

Prof. Dr. Nandamundi L. Vijaykumar (MEMBRO EXTERNO - INPE)

---

Prof. Dr. Cláudio Alex Jorge da Rocha (MEMBRO EXTERNO - IFPA)

---

Prof. Dr. Ubiratan Holanda Bezerra (MEMBRO PPGEE/UFPA)

VISTO:

---

Prof. Dr. Evaldo Gonçalves Pelaes (COORDENADOR PPGEE/UFPA)

UFPA/ITEC/PPGEE  
Campus Universitário do Guamá  
Belém – Pará - Brasil  
2016

## AGRADECIMENTOS

Agradeço primeiramente à Deus, pela minha existência e pela força que tem me dado para enfrentar as dificuldades encontradas ao longo desses anos.

Aos meus pais, pelo exemplo de determinação e garra na luta pelos nossos sonhos.

Ao meu amado esposo Marcelino Silva, pela companhia e conforto nas horas difíceis.

Aos meus amigos do LPRAD, LEA e LINC por tornarem as tarefas acadêmicas mais prazerosas e amenas.

Ao Prof<sup>o</sup> Renato Francês, um grande amigo, pelas oportunidades acadêmicas e profissional oferecidas.

Um agradecimento muito especial ao meu orientador Ádamo Santana, pela amizade e pelo apoio incondicional para finalização desse trabalho.

À UFPA, por sempre ter sido uma fonte de conhecimento e por estar sempre de portas abertas a me receber em qualquer dos seus cursos ofertados.

A todos, impossível citar todos, que de certa forma, direta ou indiretamente contribuíram para que eu chegasse até aqui.

Muito obrigada!

**Liviane Rêgo.**

## SUMÁRIO

LISTA DE ILUSTRAÇÕES .....	ix
LISTA DE TABELAS.....	x
RESUMO.....	xi
ABSTRACT .....	xii
1. INTRODUÇÃO.....	1
1.1 CONTEXTUALIZAÇÃO.....	1
1.2 DEFINIÇÃO DO PROBLEMA .....	2
1.3 METODOLOGIA .....	4
1.4 ORGANIZAÇÃO DO DOCUMENTO .....	5
2. REFERENCIAL TEÓRICO.....	6
2.1 CONSIDERAÇÕES INICIAIS .....	6
2.2 EXTRAÇÃO DE CONHECIMENTO DE BASES DE DADOS.....	6
2.3. TÉCNICAS DE CLUSTERIZAÇÃO .....	9
2.3.1 MAPAS AUTO-ORGANIZÁVEIS DE KOHONEN .....	10
2.3.2 INFORMATION THEORETIC LEARNING MEAN SHIFT .....	13
2.4 ESTRATÉGIAS DE PREDIÇÃO .....	15
2.4.1 ESTRATÉGIAS DE PREDIÇÃO DE CONSUMO .....	16
2.4.2 REDES NEURAIS ARTIFICIAIS .....	17
2.4.3 REGRESSÃO LINEAR MÚLTIPLA .....	21
2.5 CONSIDERAÇÕES FINAIS.....	23
3. TRABALHOS CORRELATOS .....	24
3.1 CONSIDERAÇÕES INICIAIS .....	24
3.2 APLICAÇÃO DE TÉCNICAS PARA PREVISÃO DE CONSUMO DE CURTO PRAZO .....	24
3.3 APLICAÇÃO DE TÉCNICAS DE CLUSTERIZAÇÃO EM ESTRATÉGIAS DE PREVISÃO .....	29
3.4 CONSIDERAÇÕES FINAIS.....	33
4. METODOLOGIA DE PREDIÇÃO DE CURTO PRAZO COM TRATAMENTO PARA OS DIAS ESPECIAIS .....	34
4.1 CONSIDERAÇÕES INICIAIS .....	34
4.2. INTRODUÇÃO DA PROBLEMÁTICA .....	34
4.2.1. PREDIÇÃO DE CONSUMO .....	36
4.2.2. DIAS ESPECIAIS E GERAÇÃO DE VALORES VIRTUAIS .....	36
4.3 DESCRIÇÃO DA METODOLOGIA PROPOSTA .....	37
4.3.1 IDENTIFICAÇÃO DA DEFASAGEM DOS DADOS .....	39
4.3.2 NORMALIZAÇÃO DOS DADOS .....	39
4.3.3 CLUSTERIZAÇÃO DOS DADOS.....	40
4.3.4 DENSIFICAÇÃO DOS DADOS .....	40
4.3.5 TRATAMENTO PARA OS DIAS ESPECIAIS.....	41
4.3.6 PREVISÃO PARA OS DIAS NORMAIS .....	41
4.3.7 PREDIÇÃO GLOBAL .....	42
4.4 CONSIDERAÇÕES FINAIS.....	42
5. APLICAÇÃO DA METODOLOGIA AO ESTUDO DE CASO DA CELPA .....	43
5.1 CONSIDERAÇÕES INICIAIS .....	43
5.2 APLICANDO A METODOLOGIA .....	43
5.2.1 ANALISANDO A BASE DE DADOS DE CONSUMO DA CELPA .....	43
5.2.2 IDENTIFICAÇÃO DA DEFASAGEM DOS DADOS .....	46
5.2.3 CLUSTERIZAÇÃO DOS DADOS.....	47
5.2.4 DENSIFICAÇÃO DOS DADOS .....	50

5.2.5 PREDIÇÃO DE CONSUMO .....	51
5.2.6 DESEMPENHO E ESCALABILIDADE .....	58
5.3 CONSIDERAÇÕES FINAIS .....	59
6. CONCLUSÕES .....	61
6.1 CONSIDERAÇÕES DA TESE .....	61
6.2 CONTRIBUIÇÕES .....	61
6.3 TRABALHOS FUTUROS E POSSÍVEIS DESDOBRAMENTOS .....	63
6.4 DIFICULDADES ENCONTRADAS .....	63
<b>BIBLIOGRAFIA .....</b>	<b>65</b>



## LISTA DE ILUSTRAÇÕES

Figura 2.1 – Fases do processo mineração de dados .....	7
Figura 2.2 – Arquitetura de uma rede de Kohonen. ....	11
Figura 2.3 – Rede de <i>Kohonen</i> . ....	12
Figura 2.4 – Representação da vizinhança ao longo do treinamento. ....	13
Figura 2.5 – Arquitetura de uma rede MLP feedforward. ....	19
Figura 2.6 – Arquitetura de uma rede MLP recorrente. ....	19
Figura 4.1 – Metodologia de Predição de Consumo. ....	38
Figura 5.1 - Série histórica de consumo da CELPA.....	44
Figura 5.2 – Padrão de consumo semanal da série.....	45
Figura 5.3 – Variações do consumo quando um feriado ocorre no meio da semana. Padrão de consumo semanal da série. ....	46
Figura 5.4 – Coeficiente de determinação $R^2$ para identificação da janela de defasagens. ....	47
Figura 5.5 – Resultados da clusterização com ITL MS. 6 clusters foram descobertos, com 12 (segunda-feira), 25 (terça-feira), 12 (quarta-feira), 23 (quinta-feira), 11 (sexta-feira santa) e 9 (sexta-feira) semanas cada. As linhas cinza representam o consumo real e a linha preta representa a moda de cada cluster. ....	48
Figura 5.6 – Resultados do mapa de Sammon.....	49
Figura 5.7 – Resultados do mapa Kohonen. 5 <i>clusters</i> foram descobertos, com 12 (segunda-feira), 26 (terça-feira), 13 (quarta-feira), 24 (quinta-feira) e 21 (sexta-feira) semanas cada. ....	50
Figura 5.8 – Séries de dados virtuais geradas pelo ITL MS após 54 iterações, com 1092 (segunda-feira), 2275 (terça-feira), 1092 (quarta-feira), 2093 (quinta-feira), 1001 (sexta-feira santa) e 818 (sexta-feira) semanas virtuais cada. As linhas cinza representam o consumo virtual e a linha preta representa a moda de cada <i>cluster</i> . ....	51
Figura 5.9 – Predição global de consumo de curto prazo.....	57
Figura 5.10 – Consumo real x consumo predito para o horizonte de médio prazo .....	58

## LISTA DE TABELAS

Tabela 3.1 – Resumo dos trabalhos correlatos .....	32
Tabela 5.1 - Melhores RNAs .....	53
Tabela 5.2 – Resultado da predição com RNA .....	53
Tabela 5.3 – RNA para predição de consumo sem densificação.....	54
Tabela 5.4 – Resumo da predição para dias normais .....	55
Tabela 5.5 – Melhor RNA .....	55
Tabela 5.6 – Resultados da predição mensal .....	58
Tabela 5.7 – Tempos de Execução .....	59

## RESUMO

Estratégias de predição de curto prazo são uma importante ferramenta usada para planejamento e operação de sistemas elétricos, bem como fundamentais para o processo de suporte à decisão para compra e venda de energia elétrica no mercado futuro. Particularmente, em se tratando de mercado de energia, uma componente importante para predição de consumo são os dias especiais (feriados ou dias atípicos, por exemplo). Tratar-se a predição de tais componentes pode ser uma tarefa complexa, dado seu comportamento atípico, quando comparado à predição de consumo em dias comuns. Em adição, via de regra, o número reduzido de amostras dificulta o treino e validação adequados dos algoritmos de predição de consumo em dias especiais. Este trabalho propõe um modelo para predição de consumo de curto prazo que utiliza a técnica *Information Theoretic Learning Mean-Shift* para clusterização e densificação dos valores de consumo em dias especiais, e algoritmos de Redes Neurais Artificiais e Regressão Linear Múltipla para predição. O modelo foi aplicado em um problema de predição de consumo da concessionária de energia elétrica da região norte do Brasil, o que proporcionou uma melhoria na acurácia dos resultados já obtido pelos métodos utilizados pela concessionária.

**Palavras-Chaves:** previsão de consumo, dias especiais, *mean shift*, clusterização e densificação.

## ABSTRACT

*The use of short-term prediction strategies is an important tool for planning and operation of electrical systems, playing a crucial part in aiding the decision support process for buying and selling of electricity in the future market. For the energy market, in particular, an important component to take into account for consumption forecasting are the special days (holidays or atypical days, for example). Given its unusual behavior, the estimation of such events can be a complex task, when compared to the forecasting of ordinary days. In addition, as they are often found with only a small number of samples, it is difficult to adequately train and validate prediction algorithms. To tackle these problems, this work presents a model for short-term load forecasting using the Information Theoretic Learning Mean-Shift model to clustering and densify the sample size of special days's events on a time series, there on followed by the prediction using statistical and/or machine learning algorithms; in this work represented by artificial neural network algorithms and multiple Linear regression. The model was applied in a load forecasting problem for the electric utility in the northern region of Brazil, providing an improvement in the accuracy of results.*

**Key Words:** *load forecasting, special days, mean shift, clustering and densify*

# 1. INTRODUÇÃO

## 1.1 CONTEXTUALIZAÇÃO

O setor elétrico tem papel estratégico e fundamental para a economia de qualquer país, dado que a energia elétrica é insumo básico de processos industriais, do comércio e da vida cotidiana das sociedades. Trata-se de um setor em que os investimentos são vultosos e demandam longo prazo para maturação e que, portanto, faz-se imperativo que as decisões a serem tomadas sejam definidas com razoável antecedência e com a máxima acurácia possível. Assim sendo, o desenvolvimento de estratégias que possibilitem a diminuição dos riscos e incertezas impostas pelo mercado e que gerem oportunidades de ganhos é fundamental para as empresas de energia.

É nesse contexto que o dimensionamento do consumo (energia elétrica requerida) futura de concessionárias do setor elétrico se torna uma importante ferramenta para o planejamento da operação elétrica, tanto nos estudos de ampliação e reforço da rede básica de geração, quanto na distribuição de energia elétrica. Objetivando melhorar tal dimensionamento, é crucial que as empresas tenham conhecimento sobre o seu mercado, pois quanto mais conhecimento acerca de tendências de mercado, mais otimizadas serão as compras de energia pelas distribuidoras e, conseqüentemente, melhor será a adequação do sistema elétrico para atender aos consumidores, quando determinadas demandas se concretizarem em um tempo futuro.

Em adição, nos últimos anos, o setor elétrico brasileiro experimentou importantes mudanças que transformaram a regulação comercial do setor em um processo bastante complexo. Mudanças essas que, impostas pela descentralização da distribuição e por um crescimento na demanda da qualidade de energia requerida pelo consumidor, resultaram no aumento pela busca de melhores sistemas de planejamento (TENFEN et al., 2013), (LINO et al., 2011).

Conforme a nova regulação comercial do setor, as empresas distribuidoras de energia devem adquirir energia para suprimento de seu mercado consumidor por meio de contratos celebrados advindos de leilões de energia específicos. Anualmente são verificadas as diferenças entre o que foi produzido ou consumido e o que foi contratado. Essas diferenças positiva ou negativa sujeitam a distribuidora ao pagamento de penalidades e, no caso da

subcontratação (compra insuficiente de energia para atendimento do mercado), existe a necessidade de compra da parte restante no mercado de curto prazo - Mercado *Spot* –, que acaba sendo um risco devido à grande volatilidade dos preços deste mercado (LUZ et al., 2012).

Considerando tal cenário, as concessionárias têm investido intensamente em técnicas que auxiliem no processo de gestão, visando a aumentar a eficiência e a confiabilidade do sistema elétrico, desde a geração até a distribuição. Em tal contexto, é imperativo o estudo de métodos de previsão de séries temporais que consigam identificar as estruturas e padrões existentes nos dados históricos e que sejam capazes de gerar os próximos padrões de consumo da série, principalmente no que se refere à previsão de consumo para uma dada região atendida por uma distribuidora de energia.

A previsão de curto prazo, que faz parte do objetivo deste estudo, é normalmente utilizada para a programação diária da operação e objetiva estabelecer os despachos eletroenergéticos otimizados de geração através da análise das condições de atendimento ao mercado de energia e demanda, considerando as condições operacionais atualizadas dos aproveitamentos hidroelétricos, das usinas térmicas, do sistema de transmissão e suas restrições elétricas. O planejamento de curto prazo tem como objetivo estabelecer diretrizes energéticas para a operação coordenada e econômica dos sistemas interligados brasileiros. As previsões podem ser revistas durante o mês, a cada semana, face às perspectivas de significativos desvios entre os valores de consumo previamente considerados em relação aos valores de consumo que vêm se configurando na operação do sistema (FREDERICO DIAS DINIZ, 2011).

## **1.2 DEFINIÇÃO DO PROBLEMA**

Em virtude da necessidade das concessionárias no sentido de melhorar sua capacidade de planejamento, aliado à disponibilidade dos recursos computacionais, as estratégias para previsão de consumo têm sido bastante utilizadas (CHO et al., 2015), (MOTLAGH et al., 2016), (MULYADI et al., 2015) e (DILHANI; JEENANUNTA, 2016), apresentando importância significativa sob a ótica financeira, contribuindo para a minimização dos custos operacionais e para a maximização da eficiência econômica das empresas.

Modelos de previsão com séries temporais continuam sendo uma das áreas investigadas pela comunidade científica, devido às suas inúmeras aplicações nos mais diversos setores de atividades acadêmicas, industriais e comerciais, conforme pode ser visto em (PAPADOPOULOS; KARAKATSANIS, 2015), (HASHMI et al., 2015) e (RAZA et al., 2014).

Uma grande variedade de técnicas tem sido desenvolvidas e empregadas com relativas vantagens e desvantagens de acordo com cada tipo de aplicação. Neste contexto, pode-se apontar a área de previsão de consumo em sistemas elétricos como um dos ramos de pesquisa largamente investigados e, por esse motivo, é utilizado como problema de aplicação da abordagem proposta neste trabalho.

Para estudos em séries temporais, uma situação ideal para a realização de previsões é o conhecimento das equações que modelam os mecanismos responsáveis pela geração de tais séries. No entanto, em muitos problemas reais essas informações não estão disponíveis, e não há condições ideais para construir equações que rejam o comportamento das variáveis de interesse. Quando isso acontece, o usual é utilizar uma abordagem baseada em modelos analíticos, na qual se tenta identificar ou aproximar o processo gerador dos dados por um modelo matemático.

Uma condição importante para se chegar a um modelo que descreva de forma satisfatória a série de dados históricas das concessionárias, obtendo-se uma previsão de consumo confiável, é a existência de uma base de dados correta e consistente. No entanto, é comum em bases de dados a presença de valores anômalos (*outliers*), que são aqueles que fogem do padrão de consumo da série. Sua ocorrência deve-se a diversos fatores, tais como: falhas na geração dos dados, comportamentos anormais no consumo (dias de feriados ou dias atípicos), entrada de novos clientes, mudanças na operação do sistema, modificações de hábitos de consumo, entre outros. A presença de dados dessa natureza degradam as séries históricas de consumo, sendo que a identificação e tratamento de tais fatores corroboram com a melhoria dos resultados de uma certa previsão.

Os valores de consumo anômalos gerados por alguns desses fatores, como é o caso da inclusão de novos clientes e a ocorrência de feriados, não podem ser caracterizados como ruídos pois possuem informações importantes e não desprezível e que deve ser incorporada ao modelo de predição. Porém, por se tratarem de eventos pontuais na série, sua ocorrência não chega a caracterizar uma tendência de comportamento na série e por isso os algoritmos de

aprendizagem de padrões tem dificuldades de descobrir seu padrão de ocorrência na série, onde alguns algoritmos chegam até a ignorar a ocorrência desses eventos na construção do modelo da série por entenderem tratar-se de um ruído. Portanto, a descoberta do padrão de comportamento desses eventos pontuais pode contribuir para o treinamento de algoritmos de predição e gerar melhores resultados de previsão. A esses eventos pontuais na série de dados que fogem do padrão da série de consumo mas que possuem informações relevantes de serem estudadas chamaremos de dias especiais.

Desta forma, com base na necessidade que as concessionárias tem de desenvolver modelos de previsão mais adequados e que melhor se ajustem aos dados históricos, é importante estudar o comportamento dos dias especiais, como o comportamento de suas ocorrências se propagam e o quanto impactam no consumo dos dias próximos a eles. Por sua vez, uma previsão com maior acurácia vem prover às concessionárias maiores ganhos financeiros em estratégias para compra de energia, como acontece no mercado *spot*, por exemplo.

### **1.3 METODOLOGIA**

Com a constatação de que a ocorrência de dias especiais impacta no histórico de consumo de energia, no atual estado da arte, uma boa alternativa para que as concessionárias planejem seus mercados seria a inclusão do conhecimento sobre esses dias em estratégias de previsão, a fim de obter cenários mais realistas e com menor erro agregado possível. Dessa forma, este trabalho propõe o desenvolvimento de uma metodologia de predição de consumo que emprega conhecimento sobre o consumo nos dias especiais a fim de melhorar a eficácia dos resultados obtidos pela previsão.

Na metodologia proposta neste trabalho, o conhecimento sobre o consumo nos dias especiais é adquirido a partir da utilização de algoritmos de clusterização e também pelo emprego da chamada densificação, que conforme a seguir, gera mais amostras de consumo diário para esses dias, chamados de valores de consumo virtuais. Dessa forma, a densificação permite aos algoritmos de descoberta de padrões aprenderem melhor o relacionamento existente entre os dados de consumo e melhorar a eficiência dos resultados da predição. A densificação é resultado da utilização do algoritmo Information Theoretic Learning Mean Shift (FUKUNAGA; HOSTETLER, 1975).

Objetivando avaliar a qualidade dos resultados obtidos pela metodologia proposta, o conhecimento adquirido com o estudo do comportamento dos dias especiais e a geração de



valores virtuais de consumo, através da clusterização e densificação, respectivamente, será aplicado em uma estratégia de predição de consumo diário para uma concessionária de energia elétrica que possui em sua série histórica valores de consumo em dias especiais.

Com base no exposto, consideram-se como objetivos específicos desta tese:

- Definição de um modelo para descoberta de padrões de consumo dos dias especiais;
- Desenvolvimento de um modelo de predição de consumo de curto prazo mais eficiente e preciso para as concessionárias de energia elétrica;
- Aplicação da densificação em uma metodologia para predição de consumo;
- Possibilidade de estender o conhecimento adquirido sobre os dias especiais para ajustes de predições para outros horizontes, como o mensal e o anual;
- Desenvolvimento de uma abordagem com alto grau de generalização, que permite realizar estudos em outros cenários do setor elétrico, tais como: operação, manutenção, supervisão e novos serviços em ambientes smart grids (e.g. smart home);

#### **1.4 ORGANIZAÇÃO DO DOCUMENTO**

Para um melhor entendimento dos temas aqui tratados, esse documento foi organizado conforme é apresentado a seguir.

No capítulo 2, são apresentados conceitos acerca das técnicas utilizadas neste trabalho, divididas entre a descoberta de conhecimento em banco de dados, bem como de técnicas de clusterização de dados e de previsão de séries temporais.

No capítulo 3, são apresentados alguns trabalhos encontrados na literatura que servem de base e ponto de apoio para as investigações propostas nesta tese.

No capítulo 4, é apresentada a metodologia proposta nesta tese a partir do detalhamento de cada uma de suas etapas.

No capítulo 5, é apresentado o estudo de caso desta tese, que se baseia na construção de um modelo de predição de curto prazo, tendo como aplicação a previsão de consumo de energia para a concessionária CELPA.

Para finalizar, no capítulo 6 são apresentadas as conclusões, descrevendo-se as contribuições, possíveis desdobramentos e dificuldades encontradas, ao longo da elaboração deste trabalho.

## 2. REFERENCIAL TEÓRICO

### 2.1 CONSIDERAÇÕES INICIAIS

Para o desenvolvimento da metodologia proposta, conforme visto no Capítulo 1, faz-se necessária a aplicação de técnicas relacionadas à extração de conhecimento em bases de dados, onde são utilizadas técnicas de clusterização, para identificação dos padrões existentes no consumo de energia, e técnicas de previsão de consumo de energia. Dessa forma, é importante descrever as técnicas de clusterização e de previsão de consumo utilizadas.

Neste capítulo, são apresentados breves conceitos sobre essas técnicas. Primeiramente, faz-se algumas considerações sobre o processo de extração de conhecimento em bases de dados, em seguida, conceitos sobre clusterização são apresentados, bem como as técnicas Mapas Auto-Organizáveis de *Kohonen* e *Information Theoretic Learning Mean-Shift*. E finalmente, na terceira seção, são exploradas definições relacionadas com estratégias para previsão, onde as técnicas Redes Neurais Artificiais e de Regressão Linear Múltipla são apresentadas.

### 2.2 EXTRAÇÃO DE CONHECIMENTO DE BASES DE DADOS

O processo de descoberta de conhecimento em bases de dados (*knowledge discovery in databases* – KDD) tem por principal objetivo encontrar padrões válidos e potencialmente úteis nos dados, que possam ser utilizados pelos usuários de níveis decisórios. O KDD desponta como uma tecnologia capaz de cooperar amplamente na busca do conhecimento embutido nos dados, de modo a apresentarem padrões que destoam, muitas vezes, do restante da massa de dados. O KDD consiste na combinação de métodos e ferramentas de estatística, inteligência artificial, visualização de banco de dados para encontrar padrões e regularidades nos dados (MANNILA, 1996).

O KDD geralmente é referenciado na literatura como *data mining* (DM) ou mineração de dados (SYMEONIDIS; MITKAS, 2006). Alguns autores consideram os termos de KDD e DM referentes a processos distintos (FAYYAD et al., 1996). Entretanto, neste trabalho, KDD e DM serão tratados de forma indistinta como o processo para extrair conhecimento a partir de base de dados.

Todo processo de DM é orientado em função de seu domínio de aplicação e dos repositórios de dados inerentes aos mesmos. Para usar os dados é necessário que estejam estruturados de forma a serem consultados e analisados adequadamente.

Existem várias formas de dividir em etapas o processo de DM, como o modelo proposto por FAYYAD et al. (1996), que dividiu o processo em nove etapas. Neste trabalho é considerada a divisão em três grandes etapas, de acordo com o modelo proposto por REZENDE (2003), conforme visto na Figura 2.1, por melhor representar o processo de desenvolvimento da metodologia proposta.

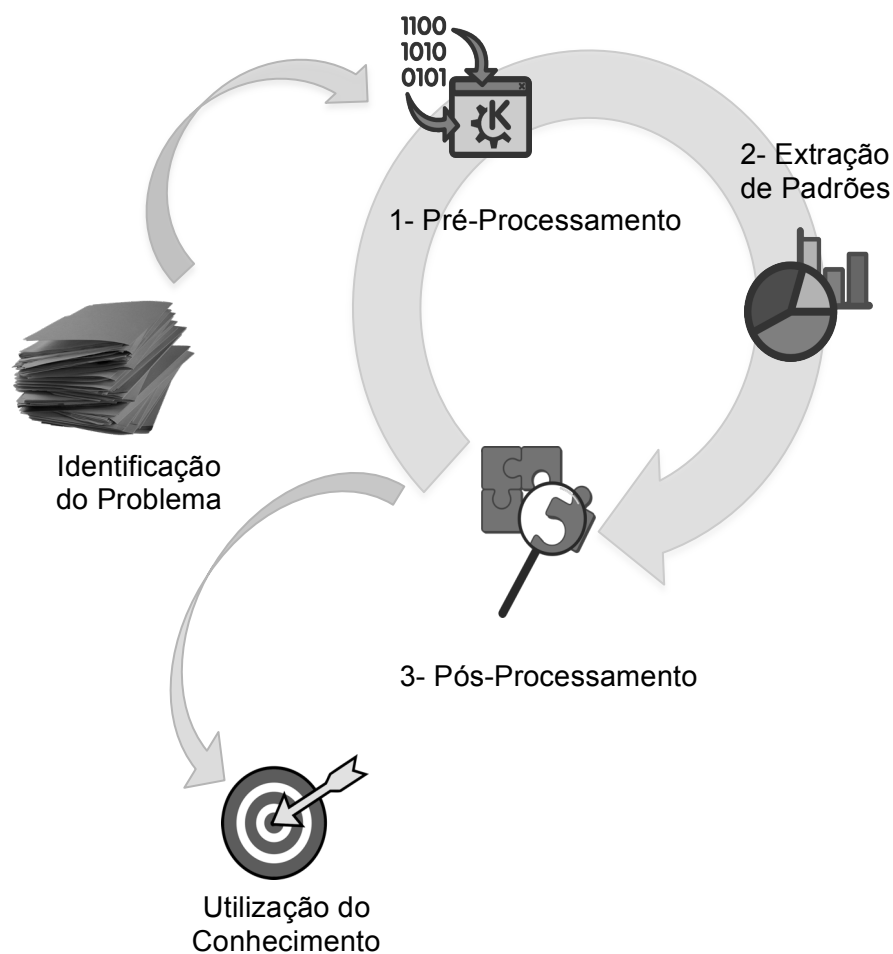


Figura 2.1 – Fases do processo mineração de dados

No modelo da Figura 2.1, foram incluídas mais duas etapas: uma fase anterior ao processo de DM, que se refere ao conhecimento do domínio e identificação do problema, e uma fase posterior ao processo, que se refere ao conhecimento obtido - essas fases não fazem parte do processo de DM mas foram incluídas na figura para representarem o contexto onde o processo de DM é aplicado. É importante ressaltar que por ser um processo iterativo, as

etapas da mineração de dados podem ser realizadas novamente, fora da sequência da figura, caso haja necessidade de ajuste em qualquer fase.

Dentro do processo DM, a primeira fase é a de Pré-Processamento, que consiste na aplicação dos métodos para tratamento, limpeza e redução do volume de dados antes de iniciar a etapa de Extração de Padrões. É importante salientar que as transformações realizadas nos dados devem ser feitas de forma criteriosa, devendo garantir que as informações presentes nos dados brutos, os padrões e tendências continuem presentes nas amostras geradas, para que os modelos finais sejam representativos da realidade inicial.

A etapa seguinte, a de Extração de Padrões, é direcionada ao cumprimento dos objetivos definidos na etapa de Identificação do Problema, onde ocorre a escolha, configuração e execução dos algoritmos para extração do conhecimento. Esta etapa se subdivide em três, a saber: (1) a escolha da tarefa de mineração de dados a ser empregada (se preditiva ou descritiva). Os algoritmos preditivos, como os de classificação e de regressão, são capazes de reconhecer a classe de novos exemplos através da generalização de exemplos passados com respostas conhecidas. E os descritivos, como as técnicas de regras de associação, clusterização e de sumarização, consistem na identificação de comportamentos intrínsecos do conjunto de dados; (2) a escolha do algoritmo (redes neurais artificiais ou redes bayesianas, por exemplo), que está subordinada à linguagem de representação dos padrões a serem encontrados; e (3) a extração dos padrões existentes, onde os algoritmos de DM escolhidos são aplicados aos dados (REZENDE, 2003).

A etapa de Extração de Padrões envolve a criação de modelos apropriados de representação dos padrões e relações identificadas a partir dos dados. O resultado desses modelos, depois de avaliados pelo analista, especialista e/ou usuário final, são empregados para prever os valores de atributos definidos pelo usuário final baseados em novos dados (FAYYAD et al., 1996).

Exemplos de padrões obtidos por meio da realização dessa etapa são equações lineares, regras, *clusters* (ou grupos) que compartilham ou apresentam padrões em comum, gráficos e estruturas, e padrões advindos de séries temporais (HAND et al., 2001). De fato, a aplicação de diversos algoritmos, modelos matemáticos e de inteligência computacional denotam o ponto crucial da etapa de Extração de Padrões.

Após a etapa de Extração de Padrões, os conhecimentos extraídos são apresentados ao usuário para interpretá-los, avaliando-os quanto à sua qualidade e/ou utilidade, para que

possam ser utilizados na tomada de decisões. Os resultados obtidos com o processo anterior podem ser avaliados conforme dois aspectos importantes, a saber: (1) quanto à sua compreensibilidade, que está relacionada com a facilidade de interpretação dos dados pelo ser humano; (2) e pela sua interessabilidade, que faz a avaliação do quanto de conhecimento interessante está presente nos dados de saída.

Após a análise do conhecimento obtido com os resultados, caso este não tenha sido satisfatório para o usuário final, o processo de Extração de Padrões pode ser repetido indefinidamente para que os parâmetros dos modelos sejam ajustados e melhores resultados sejam obtidos.

### 2.3. TÉCNICAS DE CLUSTERIZAÇÃO

As técnicas de clusterização têm sido uma boa alternativa em modelos de predição de consumo, a fim de aumentar a precisão dos resultados obtidos. Mais especificamente, essas técnicas costumam ser empregadas na fase de pré-processamento para organização dos dados de forma adequada (CHUNSHAN; XIAOFENG, 2015) e (CUGLIARI et al., 2016).

Clusterização é uma técnica usada para agrupar dados, considerando critérios de similaridade entre eles, em grupos relativamente homogêneos chamados *clusters*. Pode ser definido também como um problema de aprendizado não-supervisionado, pois a estrutura dos dados e suas propriedades, que os tornam semelhantes, são desconhecidas. Como não existem rótulos iniciais, o objetivo da clusterização é encontrar uma organização válida e conveniente dos dados, ao invés de separá-los em categorias, como acontece no reconhecimento de padrões e na classificação de dados (JAIN; DUBES, 1988).

Os elementos dentro de um *cluster* têm de ser mais semelhantes entre si do que em relação aos elementos de *clusters* distintos. Para conseguir isto, é necessário quantificar a similaridade entre os elementos, através das medidas de similaridades, que exprimem a “distância” entre dois objetos. Quanto menor o valor desta, mais semelhantes serão os objetos e deverão estes ficarem no mesmo *cluster*. De outro modo, quanto maior a “distância”, menos similares serão os objetos e, em consequência, eles deverão estar em *clusters* distintos.

Não há uma medida de similaridade que sirva para todos os tipos de variáveis que podem existir numa base de dados (COLE, 1998). Dados dois pontos em um espaço  $n$ -dimensional, a distância entre eles pode ser dada por métricas como a Euclidiana e a

Manhattan (conhecida também como *City Block*), entre outras, sendo a Euclidiana a mais utilizada (COLE, 1998) e (HAN; KAMBER, 2001).

Sendo assim, sejam  $p = (p_1, p_2, \dots, p_n)$  e  $q = (q_1, q_2, \dots, q_n)$  pontos do espaço  $n$ -dimensional. Na métrica Euclidiana, a função distância é definida como:

$$d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (2.1)$$

Na métrica de Manhattan, a função distância é definida como:

$$d(p, q) = \sum_{i=1}^n |p_i - q_i| \quad (2.2)$$

Existem diversos algoritmos de clusterização vistos na literatura. A escolha de um algoritmo depende tanto dos tipos disponíveis quanto da aplicação desejada. A classificação desses algoritmos mais bem aceita é a divisão em métodos de clusterização por particionamento e clusterização hierárquica (XU; WUNSCH, 2005). Na primeira, o conjunto de dados é dividido em um número determinado de *clusters* uma única vez. Na segunda, os dados são divididos gradualmente obtendo diversas partições dos dados.

Neste trabalho, as técnicas de clusterização empregadas foram Mapas Auto-Organizáveis de *Kohonen* e *Information Theoretic Learning Mean-Shift*, que estão descritas na sequência.

### 2.3.1 MAPAS AUTO-ORGANIZÁVEIS DE *KOHONEN*

O mapa auto-organizável de *Kohonen* é uma rede neural tipo *feedforward* não-supervisionada que representa uma técnica de mapeamento auto-organizável (*self-organizing map* - *SOM*) que possui neurônios que competem entre si para descobrir qual gera ou atribui a maior saída. A rede de *Kohonen* pode ser aplicada para problemas não-lineares de alta dimensionalidade, como extração de características e classificação de imagens e padrões, entre outras.

Os neurônios em uma rede de *Kohonen* são posteriormente ordenados e apresentados em grades (treliça), normalmente mono ou bi-dimensionais. Os neurônios se tornam seletivamente “ajustados” a vários estímulos (padrões de entrada) ou classes de padrões de entrada ao longo de um processo competitivo de aprendizado. A localização dos neurônios vencedores se torna ordenada entre si de tal forma que um sistema de coordenada significativo é criado na treliça, para diferentes características de entrada.

A arquitetura de uma rede de *Kohonen* é mostrada na Figura 2.2.

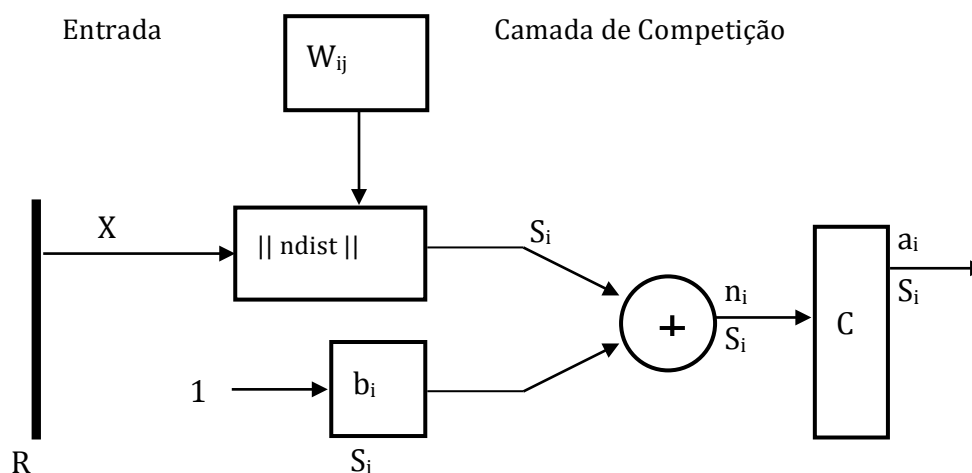


Figura 2.2 – Arquitetura de uma rede de Kohonen.

Na Figura 2.2,  $R$ ,  $X$ ,  $W_{ij}$ ,  $b_i$ ,  $S_i$ ,  $n_i$ ,  $a_i$  representam as entradas, vetor de entrada, matriz de pesos sinápticos, bias, neurônios da camada de competição, redes de entrada e a saída, respectivamente. Os elementos de  $S_i$  são resultado do cálculo da distância entre os vetores  $X$  e  $W_{ij}$ , que está representado pela caixa “ndist”.

Resumidamente, segundo *Kohonen* (1988), a rede funciona da seguinte forma: os pesos sinápticos iniciam contendo valores aleatoriamente baixos, e um sinal de entrada  $x$  (com valores que representam uma informação qualquer), com dimensão  $n$ , é provido para a rede sem que se especifique a saída desejada (característica de uma rede não-supervisionada) - o sinal de entrada  $x$  e o vetor de pesos  $W_j$  encontram-se descritos nas Equações 2.3 e 2.4. Conforme o sinal de entrada, um neurônio de saída  $y$  deverá responder melhor para a entrada respectiva, e então, será o vencedor. Isto faz com que este neurônio seja disparado sempre que essa entrada específica for apresentada à rede neural. O neurônio vencedor é aquele que possui a menor Distância Euclidiana entre o vetor de pesos por ela armazenado e o vetor de entrada, conforme Equação 2.5.

O sinal de entrada  $x$ , o vetor de pesos  $W_j$  e o cálculo do neurônio vencedor são descritos nas Equações 2.3, 2.4 e 2.5, respectivamente, a seguir:

$$x = [x_1, x_2, x_3, \dots, x_n]^{tn} \quad (2.3)$$

$$W_j = [w_{j1}, w_{j2}, \dots, w_{jm}], j = 1, 2, \dots, l \quad (2.4)$$

onde  $l$  é o número total de neurônios no mapa.

$$d_i(t) = \sum_{j=1}^N (x_j(t) - w_{i,j}(t))^2 \quad (2.5)$$

onde,  $x_j(t)$  é o padrão de entrada da rede,  $w_{i,j}(t)$  é o vetor peso do neurônio  $(i,j)$ . O índice  $(i,j)$  mostra a posição do neurônio no mapa, conforme visto na Figura 2.3.

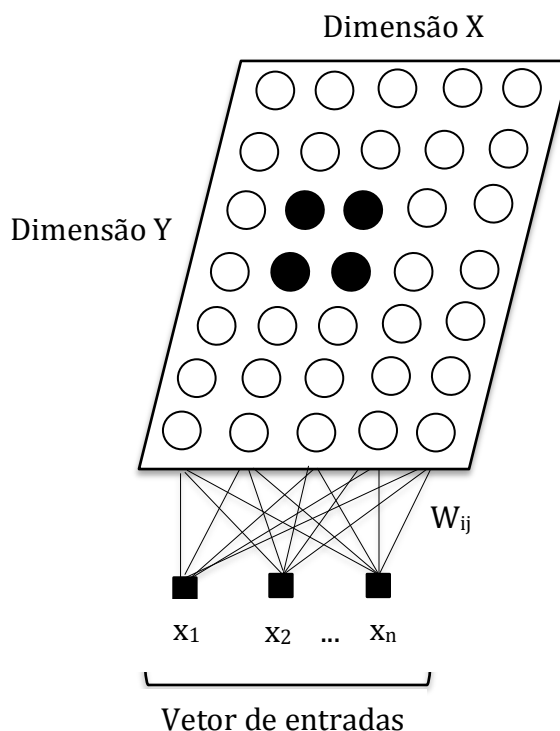


Figura 2.3 – Rede de *Kohonen*.

À medida que  $x$  é apresentado à rede, tanto o neurônio vencedor,  $y$ , quanto os seus neurônios vizinhos,  $V_y$ , terão seus pesos sinápticos ajustados para responder melhor à entrada apresentada (um neurônio ativado tende a excitar os neurônios em sua vizinhança imediata). Cada atribuição de novos valores e distâncias abrangendo toda a rede é chamada de época. Pela repetição da adaptação de pesos é possível determinar o melhor número de épocas de treinamento para cada matriz com as distâncias entre o neurônio vencedor e sua vizinhança. O neurônios nessa vizinhança são atualizados a cada iteração. Na Figura 2.4, são ilustradas a formação de vizinhança a partir do neurônio vencedor em topologia hexagonal e retangular. Algumas outras topologias que podem ser utilizadas são a triangular, quadrática e randômica.



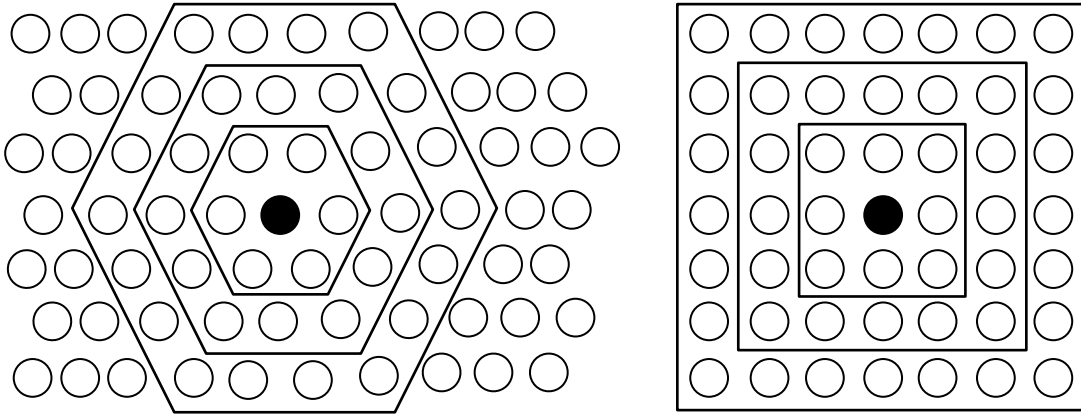


Figura 2.4 – Representação da vizinhança ao longo do treinamento.

### 2.3.2 INFORMATION THEORETIC LEARNING MEAN SHIFT (ITL MS)

O *Mean Shift* é um método não paramétrico, iterativo, com a finalidade de estimar o gradiente de uma função densidade de probabilidade, dado um conjunto de dados discretos dessa função. Este método é normalmente utilizado para encontrar modas de funções densidade, clusterização e segmentação de sinais (COMANICIU; MEER, 1999). A sua concepção se deu em 1975 por Fukunaga & Hostetler (FUKUNAGA; HOSTETLER, 1975) como um algoritmo que estima gradiente de uma função densidade com aplicações em reconhecimento de padrões. O algoritmo *Information Theoretic Learning Mean Shift* fornece meios para capturar as estruturas dominantes no conjunto de dados, através da estimação da função de densidade de probabilidade (pdf). O algoritmo ITL MS tem sido amplamente empregado em diversas áreas de processamento de imagens, como em busca de objetos, reconhecimento de cenários, entre outros (ZHOU et al., 2014), (AGARWAL et al., 2015) e (CHO; KIM, 2015). Seu uso para problemas de previsão de consumo, no entanto, é uma nova abordagem importante introduzida neste trabalho.

Considerando um conjunto de dados  $X_0 = (X_i)_{i=1}^N \in R^D$ , usando o método não paramétrico de janela de parzen (PARZEN, 1962), a pdf pode ser estimada por:

$$p(x, \sigma) = \frac{1}{N} \sum_{i=1}^N G_{\sigma}(x - x_i) \quad (2.6)$$

onde  $G_{\sigma}$  é um kernel gaussiano dado por  $G(t) = e^{-\frac{1}{2}t^2}$  com largura de banda  $\sigma > 0$ .

O objetivo do algoritmo é encontrar as modas da base de dados em que  $\nabla p(x) = 0$ . A equação do ponto estacionário, é:

$$m(x) = \frac{\sum_{i=1}^N G_{\sigma}(x-x_i) \cdot x_i}{\sum_{i=1}^N G_{\sigma}(x-x_i)} \quad (2.7)$$

A diferença  $m(x) - x$  é conhecida como mean shift.

O ITL MS segue o critério de utilização que minimiza a entropia de  $X$ ,  $H(X)$ , enquanto mantém a distância de Cauchy-Schwartz  $D_{CS}(X, X_0)$ , entre  $X$  e  $X_0$  fixado em algum valor de  $k$  (RAO et al., 2006).

A entropia quadrática de Renyi's (RÉNYI, 1976) para uma pdf pode ser calculada através da fórmula:

$$H(X) = -\log \int_{-\infty}^{+\infty} p^2(x) dx \quad (2.8)$$

Substituindo (2.6) em (2.8), temos:

$$H(X) = -\log V(X) \quad (2.9)$$

com

$$V(X) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N G_{\sigma'}(x_i - x_j) \quad (2.10)$$

onde  $\sigma' = \sqrt{2\sigma}$ .  $V(X)$  é conhecido como a informação potencial da pdf  $p(X)$ .

Para medir a entropia cruzada entre duas pdfs, temos:

$$H(X, X_0) = -\log V(X, X_0) \quad (2.11)$$

com

$$V(X, X_0) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N G_{\sigma'}(x_i - x_{0j}) \quad (2.12)$$

A distância Cauchy-Schwartz entre duas pdfs  $p$  e  $q$  podem ser calculadas usando:

$$D(X, X_0) = \log \frac{(\int p^2(x) dx)(\int q^2(x) dx)}{(\int p(x)q(x) dx)^2} \quad (2.13)$$

Utilizando o conceito de que o cosseno de um ângulo entre dois vetores é a medida da distância entre eles, temos:

$$D_{CS}(X, X_0) = -[H(X) + H(X_0) - 2H(X, X_0)] \quad (2.14)$$

A fórmula de otimização irrestrita sob o parâmetro  $\lambda$ , que representa o deslocamento entre os dois objetivos, é:

$$J(X) = \min H(X) + \lambda[D_{CS}(X, X_0) - k] \quad (2.15)$$

Diferenciando  $J(X)$  em relação a cada  $x_i \in X$  permite a transformação de  $X_0$  em outro conjunto  $X^{t+1}$  na iteração  $t+1$ , que faz o uso da informação contidos na pdf de  $X$  na iteração  $t$  (RAO et al., 2006).

O agrupamento das observações em diferentes classes é feito através do deslocamento de cada observação por uma certa quantidade proporcional ao gradiente no ponto de observação (FUKUNAGA; HOSTETLER, 1975). Além disso, cada geração (deslocamento) de pontos  $x_i^t \in X^t$  descreve uma pdf de  $X^t$  que retém informações de pdf de  $X_0$ . O conjunto  $X_V = \{X^1 \cup X^2 \dots \cup X^t\}$  representa os dados virtuais gerados pelo algoritmo ITL MS; formando um *cluster* denso que possui propriedades compatíveis com a série original  $X_0$ . Esta propriedade é usada neste trabalho, de uma nova maneira, para densificar os *clusters* através da geração de pontos ao longo das iterações do algoritmo e usá-los como novos pontos da base dados (dados virtuais), que são compatíveis com a pdf original; a geração de  $X_V$  e o seu uso em algoritmos de predição é chamado de densificação.

## 2.4 ESTRATÉGIAS DE PREDIÇÃO

Previsões são insumos importantes para realizar planejamentos em geral. Finanças precisam de previsões para projetar fluxos de caixa e necessidades de capital. Recursos humanos precisam de previsões para prever as necessidades de contratação e treinamentos. No setor elétrico, as estratégias de previsão são utilizadas para adequação do sistema energético à demanda de consumo, em tarefas de planejamento de mercado e de operacionalização do sistema.

Criar cenários através da realização de projeções é, independente do domínio de aplicação, uma alternativa eficiente que visa garantir a operação de forma satisfatória de sistemas em geral. Nesse sentido, para obter e confirmar uma demanda futura, facilitando a programação de recursos e garantindo o ganho de uma oportunidade de mercado, por exemplo, faz-se necessário a utilização de métodos estatísticos ou computacionais, regressão linear e análise de correlação, médias móveis, método sazonal multiplicativo, redes neurais artificiais, entre outros. A escolha por um ou outro método depende especificamente da análise dos dados históricos e de suas especificidades, onde a acurácia desejada e complexidade computacional são fatores importantes nessa decisão.

### 2.4.1 ESTRATÉGIAS DE PREDIÇÃO DE CONSUMO

Objetivando otimizar o planejamento e a operação dos sistemas elétricos, reduzindo custos e aumentando a qualidade do fornecimento de energia elétrica, a previsão de consumo é tida como uma estratégia importante para o setor energético pois é a partir dela que se planejam e operam esses sistemas de forma confiável e segura (QUILUMBA, 2015) e (RAZA et al., 2014). Através da previsão de consumo é que o sistema elétrico se adequa para atender os consumidores, quando determinadas demandas se concretizarem em um tempo futuro, proporcionando uma diminuição da diferença entre a quantidade de energia comprada e consumida.

Para a realização de previsões é necessário o conhecimento das equações que modelam os mecanismos responsáveis pela geração das séries temporais. No entanto, em muitos problemas reais essas informações não estão disponíveis, e não se tem condições ideais para construir equações que descrevam o comportamento das variáveis de interesse. Quando isso acontece, utiliza-se uma abordagem baseada em modelos, na qual se tenta identificar ou aproximar o processo gerador dos dados. O ideal é obter o processo gerador dos dados com o menor erro possível, evitando assim gastos desnecessários ou má qualidade do fornecimento de energia elétrica. Um alto grau de precisão possibilita a otimização da produção, redução dos custos de produção e melhor programação de intercâmbio de energia entre concessionárias (HONG; SHU, 2016).

A previsão de consumo pode ser dividida, de acordo com o horizonte de previsão, em três categorias: de curto prazo (horizonte geralmente de um hora a uma semana), médio prazo (previsões de uma semana a poucos meses) e de longo prazo (horizontes de previsão maiores que um ano). Os diferentes horizontes de previsão são importantes para diferentes aplicações em uma concessionária de energia elétrica: a maioria das previsões de longo prazo tem sido utilizada com o propósito do planejamento dos sistemas de transmissão e distribuição. Já as previsões de médio e curto prazo são constantemente utilizados no processo de compra de energia e para gerenciamento pelo lado da demanda (*Demand-Side Management - DSM*) (GUERINI; DE NICOLAO, 2016).

Vários métodos para prever consumo de energia elétrica têm sido propostos, na tentativa de diminuir os riscos de operação de sistemas elétricos de potência, sendo que o desempenho dos modelos varia de acordo com cada sistema, com os dados utilizados, com o horizonte de previsão, entre outros. Esses métodos podem ser divididos em dois tipos:

técnicas tradicionais (regressão, séries temporais, etc.) e técnicas de inteligência artificial. A seguir, serão detalhadas uma técnica de cada tipo: primeiro, redes neurais artificiais e em seguida, regressão linear múltipla, que foram as técnicas utilizadas na metodologia proposta neste trabalho.

#### 2.4.2 REDES NEURAIIS ARTIFICIAIS

Uma rede neural artificial (RNA) é um modelo computacional, inspirado biologicamente no funcionamento/atividade do cérebro, formado por elementos de processamento (chamados de neurônio) e conexões entre esses elementos, com pesos ligados a essas conexões, formando assim uma estrutura neural. Sobre essa estrutura, são implementados algoritmos de aprendizado que ajustam os parâmetros da estrutura de acordo com o tipo de conhecimento a ser codificado na rede. RNAs são chamadas de modelos conexionistas devido à grande importância das conexões entre os neurônios para o processamento da rede pois são os pesos das conexões os responsáveis pelo “conhecimento” codificado nas RNAs.

Os modelos de redes neurais têm sido utilizados como alternativas aos modelos tradicionais de predição, classificação ou aglomeração por se caracterizarem como aproximadores universais de funções, segundo HAYKIN (2004), e devido às suas características de desempenho frente a dados incompletos e/ou sua capacidade de estabelecer relações entre os mesmos.

As RNAs possuem algumas características que as tornam alvos de intensas pesquisas, tais como:

- **Aprendizado e Adaptação:** uma das propriedades mais importantes de uma RNA é a capacidade de aprender por intermédio de exemplos e realizar inferências sobre o que aprendeu, melhorando gradativamente o seu desempenho;
- **Generalização:** RNAs são capazes de generalizar o seu conhecimento a partir de exemplos anteriores e com isso lidar com informações nunca antes vista no conjunto de treinamento;
- **Processamento Paralelo:** característica de processamento intrínseca das RNAs e que foi herdada da sua inspiração biológica, o cérebro;

- **Robustez:** a perda de um conjunto de elementos processadores (neurônios) pode não causar o mau funcionamento da RNA.

Existe uma variedade bastante razoável de modelos de RNAs que executam vários tipos de tarefas. A tarefa de previsão de séries temporais envolve a construção de modelos a partir de dados históricos em um determinado período. Assim, o emprego de RNAs na previsão de séries temporais tem sido bastante utilizado pela sua capacidade de adaptação ao ambiente de previsão e, assim, modelar o comportamento da série através do aprendizado dos dados históricos; provendo assim um mapeamento entre as variáveis de entrada e a previsão na saída da rede (VELASCO et al., 2015). Outro fator favorável às RNAs está relacionado ao fato delas possuírem a habilidade de aproximar qualquer função contínua usando algoritmos de aprendizado através de uma representação interna sem a necessidade de construção de um modelo matemático explícito que relaciona suas entradas e saídas (SULAIMAN et al., 2016).

Uma característica importante para a construção de uma RNA é a definição de sua arquitetura. Na seção a seguir encontra-se descrita a arquitetura de uma rede Perceptron Multi-camadas, uma das mais utilizadas na literatura.

#### 2.4.2.1 PERCEPTRON MULTICAMADAS

A arquitetura da rede Perceptron Multi-camadas (*MultiLayer Perceptron – MLP*) (Figura 2.5) é composta por neurônios conectados, seguindo uma formação em camadas. Essas conexões são responsáveis por propagar as entradas da RNA. As unidades de entrada, que compõem essa camada, têm o objetivo de difundir o sinal inicial sem nenhuma modificação para a segunda camada. Os dados são apresentados à rede pela camada de entrada, são processados pelas camadas subsequentes e por último a rede gera uma saída para a informação que lhe foi apresentada. É importante salientar que se cada camada de neurônios é somente ligada à camada subsequente (menos a camada de saída, que só recebe ligações) e a informação trafegar em um único sentido, a rede é chamada de *feedforward*, Figura 2.5. Caso existam conexões de retorno entre as camadas, posterior para a anterior, a rede é chamada de recorrente, Figura 2.6.

Redes MLP apresentam um poder computacional maior do que aquele apresentado pelas redes perceptron que possuem uma única camada (HAYKIN, 2004), (ZHANG et. al., 1998) e (KASABOV, 1996). Tal poder computacional é conseguido com a adição de uma ou mais camadas intermediárias entre a entrada e a camada de saída. Em CYBENKO (1988), foi

provado que são necessárias, no máximo, duas camadas intermediárias, com um número suficiente de unidades por camada que é definido de forma empírica, para se aproximar qualquer função, seja ela linearmente separável ou não. Também foi provado que apenas uma camada intermediária é suficiente para aproximar qualquer função contínua.

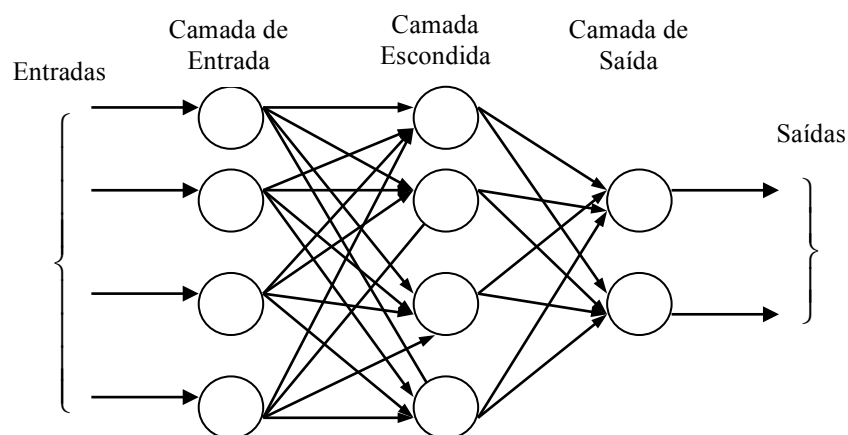


Figura 2.5 – Arquitetura de uma rede MLP feedforward.

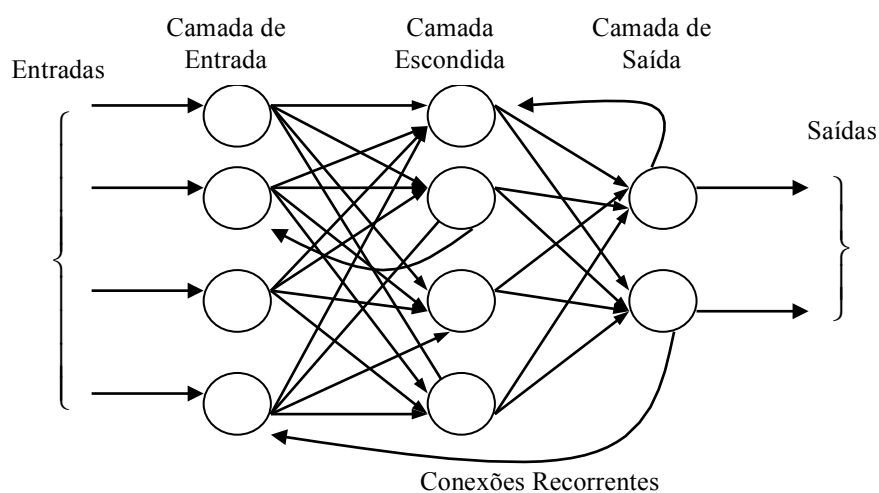


Figura 2.6 – Arquitetura de uma rede MLP recorrente.

Assim, a definição da arquitetura de uma rede MLP apresenta três características distintas:

- **Definição da Estrutura:** determinação do número de camadas intermediárias e das quantidades de neurônios em cada uma dessas camadas;

- **Tipo de Conexão:** determinação do tipo de conexões entre as camadas da rede (*feedforward* ou recorrente);
- **Função de Ativação:** determinação das funções de ativação dos neurônios. Os neurônios da mesma camada apresentam os mesmos tipos de funções de ativação, enquanto que neurônios de camadas diferentes podem ou não, dependendo do problema de aplicação, apresentar os mesmos tipos dessas funções.

#### 2.4.2.2 ALGORITMOS DE TREINAMENTO

Existem inúmeros algoritmos para o treinamento de redes neurais, sendo que os mais conhecidos são os algoritmos *Backpropagation* e o *Levenberg-Marquardt*. O algoritmo de *Backpropagation* é um dos mais importantes no desenvolvimento das redes neurais, no entanto sua taxa de convergência é muito baixa. Por isso, foram propostos outros algoritmos como alternativa ao *backpropagation* (ANDERSEN; MILAMOWSKI, 1995), (BATTITI, 1992) e (HAGAN; MENHAJ, 1994), sendo que o algoritmo de *Levenberg-Marquardt* (LM) tem sido amplamente utilizado e aceito como algoritmo muito eficiente para solucionar o problema de minimização de quadrados para funções não-lineares.

Enquanto o *Backpropagation* padrão utiliza a descida de gradiente como método de aproximação do mínimo da função erro, o algoritmo de *Levenberg-Marquardt* utiliza uma aproximação pelo método de Newton (HAGAN; MENHAJ, 1994). Esta aproximação é obtida a partir da modificação do método de Gauss-Newton, introduzindo-se o parâmetro  $\mu$ , conforme a equação:

$$\Delta x = [J^T(x)J(x) + \mu I]^{-1}J^T(x)e(x) \quad (2.16)$$

onde  $I$  é a matriz identidade,  $e(x)$  é o erro e  $J$  é a matriz Jacobiana. O parâmetro  $\mu$  é multiplicado por um fator  $\beta$  toda vez que um passo resultar num aumento na função erro, que se deseja minimizar. Quando um passo resultar na diminuição da função genérica  $V(x)$ ,  $\mu$  é dividido pelo fator  $\beta$ . Isto significa dizer que, se houver uma convergência para o mínimo da função,  $\mu$  é pequeno e o algoritmo se aproxima do método de Newton (passo  $I/\mu$ ); no caso de não haver convergência, o método se aproxima da descida de gradiente. O parâmetro  $\mu$  funciona como um fator de estabilização do treinamento, ajustando a aproximação de forma a utilizar a rápida convergência do método de Newton e evitando passos muito grandes que possam levar a um erro de convergência (HAGAN; MENHAJ, 1994).



Esse método apresenta convergência em menos iterações, mas requer mais cálculos por iteração, devido ao cálculo de matrizes inversas. Apesar do grande esforço computacional, ele segue sendo o algoritmo de treinamento mais rápido para redes neurais, quando se trabalha com um número moderado de parâmetros na rede. Se esse número é elevado, a utilização desse algoritmo é pouco prática.

### 2.4.3 REGRESSÃO LINEAR MÚLTIPLA

A regressão é uma técnica estatística multivariada para previsão de valores de uma variável de resposta (dependente) através de uma variável explicativa (independente) – regressão simples – ou através de diversas variáveis independentes (regressão múltipla).

Os modelos de análise multivariada podem ser classificados de acordo com o número de variáveis trabalhadas (RENCHEER, 2002):

- Regressão simples: com apenas uma variável Y (dependente) e uma variável X (independente);
- Regressão múltipla: uma variável Y e várias variáveis X;
- Regressão múltipla multivariada: atendendo várias variáveis Y e X.

A técnica de regressão linear múltipla (RLM) denota um modelo específico de análise multivariada cuja finalidade principal é obter uma relação matemática entre uma das variáveis (a variável dependente) e outras variáveis que descrevem o domínio (variáveis independentes), estabelecendo o grau de explicabilidade que estas possuem para com a variável dependente; predizendo o valor de uma única variável com base em várias outras variáveis preditoras (HARRIS, 2001). Quanto mais significativo for o peso de uma variável isolada, ou de um conjunto de variáveis explicativas, tanto mais se poderá afirmar que alguns fatores afetam mais o comportamento de uma variável de resposta especificamente procurada, do que outros, podendo dessa forma, ser utilizada também para a avaliação dos efeitos das variáveis explicativas como previsoras das variáveis de resposta.

Para uma amostra aleatória de tamanho n, o modelo de regressão múltipla é dado pela seguinte equação (LEMOS; MYNBAEV, 2004):

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i \quad (2.17)$$

onde,  $Y_i$  é a variável alvo,  $\beta_0$  é o intercepto do eixo y,  $X_i$  são as variáveis independentes, os demais valores de  $\beta_i$  são os coeficientes angulares da i-ésima variável, k é o número de

variáveis independentes e  $u_i$  são os erros aleatórios. A letra  $i$  refere-se a  $i$ -ésima observação de uma amostra.

O sistema geral de regressão múltipla pode ser então representado da seguinte forma:

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_k X_{1k} + u_1 \\ Y_2 &= \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_k X_{2k} + u_2 \\ &\vdots \\ Y_n &= \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_k X_{nk} + u_n \end{aligned} \quad (2.18)$$

Este sistema de equações pode ser representado como um sistema matricial.

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix} \times \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad (2.19)$$

Essa especificação objetiva gerar os valores dos parâmetros do vetor  $\beta$ . Para isto, emprega-se o método dos mínimos quadrados ordinários (MQO). Ele permite gerar estimativas para os parâmetros  $\beta$  com menor variância e não-viesados. O método de MQO consiste em minimizar a soma dos erros quadrados da regressão estimada (Equação 2.20), de tal forma que este seja o menor possível.

$$\min \sum_i^n u_i^2 = \sum_i^n (Y_i - \hat{Y}_i)^2 \quad (2.20)$$

O resultado da aplicação do método de MQO ao modelo (2.19) é dado por:

$$A = (X^t X)^{-1} \times X^t Y \quad (2.21)$$

A regressão linear múltipla pode ser utilizada, entre outras possibilidades, para:

- Determinar se existe uma relação entre a variável dependente e a(s) independente(s);
- Avaliar a intensidade da relação;
- Determinar a estrutura ou a forma da relação (equação);
- Predizer os valores da variável dependente.

## 2.5 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados conceitos acerca de descoberta de conhecimento em banco de dados, clusterização e predição, que são processos utilizados no modelo de predição de consumo proposta neste trabalho.

Foram descritas também técnicas que podem ser empregadas nessa metodologia, como as técnicas ITL *Mean-Shift* e redes de *Kohonen*, utilizadas como técnicas de clusterização e as técnicas RNA e RLM, utilizadas em estratégias de predição.

No próximo capítulo, será apresentado o estudo dos trabalhos correlatos que realiza um apanhado geral, para efeito de comparação com a proposta desta tese, sobre os trabalhos que vem sendo publicados na literatura especializada.

### 3. TRABALHOS CORRELATOS

#### 3.1 CONSIDERAÇÕES INICIAIS

Neste capítulo, é realizado o levantamento do estado-da-arte de modelos que vêm sendo utilizados para realizar predição de consumo de energia elétrica de curto prazo, associados com estratégias para identificação e tratamento de *outliers* - dias especiais também - presentes em bases de dados de consumo.

Com as referências apresentadas, ratifica-se que pesquisas relacionadas com a problemática apontada nesta tese, como a predição de consumo de energia de curto prazo e estudo e tratamento de *outliers*, representam uma área de pesquisa ativa e em aberto na literatura.

A apresentação dos trabalhos correlatos está organizada em duas categorias, a saber: aplicação de técnicas para previsão de consumo de energia de curto prazo e estudos da aplicação de técnicas de clusterização em estratégias de previsão.

#### 3.2 APLICAÇÃO DE TÉCNICAS PARA PREVISÃO DE CONSUMO DE ENERGIA DE CURTO PRAZO

Previsão de consumo é um importante instrumento em modernos sistemas de energia, bastante utilizada na operacionalização dos sistemas, no desenvolvimento de estratégias de fornecimento de energia, no planejamento financeiro e gerenciamento do mercado elétrico (QUILUMBA, 2015) e (RAZA et al., 2014). Com essa finalidade, métodos estatísticos e de inteligência computacional têm sido amplamente empregados (PAPARODITIS; THEOFANIS, 2013), (HONG; SHU, 2016) e (MOTLAGH et al., 2016).

Uma tarefa que busca melhoria dos resultados em problemas de predição é descobrir como o consumo se relaciona com outras variáveis. Desta forma, diversos modelos que empregam variáveis exógenas ao sistema elétrico em suas estratégias de previsão vem sendo empregados na literatura. É o que pode ser constatado com as referências (SHAO et al., 2015), (GARULLI et al., 2015) e (PAPADOPOULOS; KARAKATSANIS, 2015), descritas a seguir.

Em (SHAO et al., 2015) é proposto um método semi-paramétrico para predição de consumo de energia de médio prazo que busca identificar fatores que impactam significativamente no consumo de energia através da utilização da métrica chamada divergência Kullback-Liebler (K-L), uma medida de informação que se baseia em medidas de similaridades. Um estudo de caso relativo à previsão de demanda de energia na China demonstrou a aplicabilidade da abordagem proposta e verificou uma dependência funcional explícita entre variáveis exógenas ao sistema elétrico e o consumo de energia. Apesar da complexidade do método, houve uma redução no erro de previsão devido à utilização de três indicadores que influenciaram no consumo: depósitos em instituições financeiras, exportações e importações.

Outro exemplo que utiliza conhecimentos exógenos aos sistemas elétricos em predição foi descrito em GARULLI et al. (2015), onde foi desenvolvido um modelo de predição baseado em testes de caixa preta e de caixa-cinza utilizando modelos auto regressivos. Visando o aprendizado de estruturas não lineares, o trabalho emprega técnicas de *Support Vector Machine* (SVM) e de redes neurais. Esses algoritmos utilizam como variáveis de entrada, entre outras, os valores de demanda ativa (DA) da rede e valores de temperatura diária (real e previsto). Objetivando analisar como a demanda ativa impacta na previsão de consumo, os modelos de predição foram testados com as mesmas entradas, exceto a demanda ativa. Os experimentos mostraram que os modelos lineares tiveram excelentes resultados nas predições para um dia à frente. Já nas previsões dos componentes de baixa frequência os modelos de SVM e de redes neurais tiveram melhores resultados. Os modelos que ignoraram a presença explícita da DA tiveram resultados de predição de consumo insatisfatórios, mostrando assim a importância dessa variável em estratégias de predição diária.

Em (PAPADOPOULOS; KARAKATSANIS, 2015), o objetivo é prever o consumo para um dia à frente em Nova Inglaterra. Para isso, quatro métodos diferentes são discutidos e comparados, são eles: SARIMA (integrado autorregressivo e médias móveis com sazonalidade), SARIMAX (integrado autorregressivo e médias móveis com sazonalidade com variáveis exógenas), *random forests* (RF) e o gradiente *boosting regression trees* (GBRT). Visando aumentar a acurácia da predição dos modelos, a temperatura foi considerada como variável de entrada, dado seu alto grau de correlação com o consumo. Os resultados deste estudo mostraram que o modelo GBRT é superior aos outros em previsões, pois obteve os menores erros.

Uma questão relacionada com a problemática de predição de consumo que vem sendo abordada na literatura, é a presença de dados anômalos (*outliers*) nas bases históricas de consumo. Nesses sistemas, consideram-se *outliers* a ocorrência de eventos, como: falhas na geração dos dados, comportamentos anômalos no consumo (dias de feriados ou dias atípicos) e modificações de hábitos de consumo (MULYADI et al., 2015). Conforme (FIDALGO; LOPES, 2005), a presença de anomalias nos históricos de dados são fatores que degradam a qualidade das predições de consumo e sua identificação e tratamento corroboram com a melhoria dos resultados de uma previsão. (LAHOUAR; SLAMA, 2015) afirma que é muito difícil prever consumos de feriados, bem como de dias que antecedem e sucedem os mesmos, uma vez que eles têm um padrão único de consumo e pelo seu número limitado de ocorrências - cada feriado ocorre uma vez por ano e o número de anos históricos que pode ser efetivamente usado para a previsão de consumo é limitada.

A previsão de consumo diária para os feriados (conhecido também como dias especiais) é uma importante estratégia explorada na literatura, alguns exemplos de publicações são (LUTHULI; FOLLY, 2016), (MULYADI et al., 2015), (RAZA et al., 2014), (ARORA; TAYLOR, 2013) e (SAIPUNYA et al., 2014). Em alguns casos, quando se trata da previsão de um dia especial, o especialista faz ajustes manuais no valor do consumo predito (LAHOUAR; SLAMA, 2015). Outros trabalhos propõem modelos que utilizam algum algoritmo que gera automaticamente a classificação do tipo do dia analisado, como o utilizado por (PANAPAKIDIS; PAPAGIANNIS, 2014). Em algumas situações, os autores não fazem nenhum tratamento em cima dos feriados, mas em virtude disso os erros encontrados para os dias especiais (finais de semana e feriados) não são tão bons quando comparados com os resultados para dias normais, como encontrado em HASHMI et al. (2015).

Seguindo a ideia de diferenciar dias normais (dias de trabalho) de dias não trabalhados (finais de semanas e feriados), em RAZA et al. (2014) é realizada previsão de consumo de curto prazo utilizando Otimização por Enxame de Partículas (*particle swarm optimization* ou PSO) baseada em RNA (PSONN). Nesse trabalho, é analisado o impacto que as informações sobre dia da semana, a hora do dia e o tipo de dia (se é um dia de trabalho ou não) tem sobre a predição de consumo de energia. A informação sobre o tipo de dia é considerado importante pelo autor, dado que o padrão de consumo de um dia trabalhado é diferente do padrão de consumo em um final de semana ou feriado, devido a mudanças nas atividades humanas. Os resultados obtidos com modelo da PSONN são comparados com outro modelo que utiliza

*Levenberg Marquardt* (LM), baseado em RNA (LMNN), onde o modelo da PSONN gerou melhores resultados.

Além da identificação se o dia é especial ou não, outra forma de utilizar o conhecimento acerca dos feriados em estratégias de predição é descobrir padrões comuns existentes em determinados grupos de feriados. Em (ARORA; TAYLOR, 2013), é proposto um modelo de predição baseado em regras que identificam na série histórica de consumo os dias que possuem perfil mais adequados para estimar a energia a ser consumida no futuro para um dia especial (feriados). A formulação dessas regras foi realizada de forma subjetiva através de inferência direta nos dados, onde cada tipo de feriado possui uma regra específica. O artigo propõem a geração de quatro regras específicas: (1) para os feriados no qual seu histórico sempre ocorre no mesmo dia do ano anterior (sexta-feira santa, por exemplo); (2) inclui os dias especiais que são do mesmo tipo e que ocorrem no mesmo dia da semana (ano novo dos anos de 2008 e de 2002, por exemplo); (3) agrupa os feriados conforme sua ocorrência, se é um dia de semana ou final de semana; (4) diferencia os feriados que ocorrem de terça a quinta-feira de outros feriados que ocorrem de sexta a segunda-feira. Por fim, regras análogas às geradas foram utilizadas com modelos de previsão baseados em RNA, decomposição de valores singulares (*singular value decomposition* ou SVD), métodos de amortecimento exponencial HWT e ARMA.

Outro exemplo de como a diferenciação dos dias pode ser utilizado em estratégias de predição de curto prazo, é o desenvolvido em SAIPUNYA et al. (2014), onde oito modelos diferentes de Máquinas de Vetor de Suporte para Regressão (VSR) foram implementados para prever o consumo diário dos diferentes dias da semana (segunda a domingo) e feriados. Neste artigo, como variáveis de entrada foram utilizados os valores de consumo horário e a temperatura.

Nos artigos apresentados nesta seção, que norteiam o arcabouço encontrado na literatura, observa-se a tendência em realizar predição de consumo de curto prazo utilizando informações exógenas ao sistema elétrico, conforme visto em (SHAO et al., 2015), (GARULLI et al., 2015), (SAIPUNYA et al., 2014) e (PAPADOPOULOS; KARAKATSANIS, 2015). Apesar dessa ser uma abordagem largamente utilizada para prever o consumo, a utilização de variáveis exógenas pode tornar as estratégias de predição impraticáveis, devido ao fato de que tais variáveis podem ser de difícil obtenção e/ou podem possuir um custo elevado. O uso de alguns atributos, como os dados climáticos e socioeconômicos, também é descartado para a análise preditiva implementada nesta tese

porque além de não estarem facilmente disponíveis e muitas das vezes serem pagos, esses dados precisam também ser estimados. Sua estimação, que não é uma tarefa simples, pode inserir erros no modelo que vão sendo acumulados, principalmente quando o período de previsão aumenta. Uma abordagem interessante, dada a importância das variáveis exógenas sobre o consumo, seria sua utilização apenas como suporte ao modelo de predição, ou seja, adicionando algum conhecimento sobre o consumo que poderia ajustar o algoritmo de predição.

Finalmente, em virtude dos pontos elencados para a previsão de consumo, na metodologia proposta nesta tese pretende-se obter melhores resultados de previsão utilizando-se apenas informações intrínsecas ao sistema elétrico. Essa característica pode ser considerada um diferencial da metodologia, pois favorece a construção de um modelo mais simples com variáveis de mais fácil obtenção pela concessionária.

Outro diferencial encontrado entre o modelo desta tese e os artigos apresentados está relacionada com a identificação dos padrões de consumo dos dias especiais. Os artigos descritos aqui, conforme vistos em (RAZA et al., 2014), (LAHOUAR; SLAMA, 2015), e (SAIPUNYA et al., 2014), não apresentam estratégias automáticas para descobertas dos padrões de consumo para esses dias, apresentando, ao invés disso, formas subjetivas (inferenciadas pelo próprio analista) ou exatas (classificando apenas em feriados ou não, ou considerando o comportamento de um dia especial como um outro dia da semana). Um exemplo de subjetividade é o modelo descrito em (ARORA; TAYLOR, 2013), onde a classificação foi feita de forma particular pelo analista. No modelo implementado nesta tese, empregam-se técnicas de clusterização para classificação dos dados (se especiais ou não), o que se caracteriza como um diferencial, pois possibilita a descoberta de padrões passíveis de serem desconsiderados em uma análise subjetiva ou exata feita por analistas.

Outra problemática, que o modelo objeto desta proposta de tese aborda, está relacionada com a presença de dados anômalos, ausentes ou insuficientes na série histórica de consumo, conforme visto em (HASHMI et al., 2015). A quantidade de valores insuficientes em bases de dados de predição impacta diretamente na aplicação dos algoritmos de predição, pois impossibilita esses algoritmos descobrirem e/ou aprenderem o comportamento da série de dados, gerando valores de erros mais elevados. Nesse sentido, a fim de evitar tais problemas, a metodologia proposta nesta tese utiliza a estratégia para geração de novas amostras para históricos de consumo que possuem valores insuficientes para utilização em algoritmos de predição, o que pode ser considerada outra importante contribuição deste trabalho.



### 3.3 APLICAÇÃO DE TÉCNICAS DE CLUSTERIZAÇÃO EM ESTRATÉGIAS DE PREVISÃO

Em modelos de predição de consumo de energia, a fim de melhorar o desempenho de seus resultados, é fundamental na fase de pré-processamento organizar os dados de forma adequada, agrupando-os de acordo com perfis semelhantes encontrados. Nesse sentido, em diversos trabalhos encontrados na literatura especializada tem-se utilizado técnicas de clusterização como uma boa alternativa, conforme visto em (PANAPAKIDIS; PAPAGIANNIS, 2014), (PANAPAKIDIS et al., 2015), (WANG, 2015), (CHUNSHAN; XIAOFENG, 2015) e (CUGLIARI et al., 2016). Na maioria dos trabalhos, um dos objetivos de utilizar a clusterização antes da etapa de predição é a identificação de fatores que podem influenciar no consumo de energia, melhorando assim, a acurácia dos resultados. Outro objetivo encontrado para clusterização na etapa de pré-processamento é de organizar os dados, agrupando-os conforme padrões semelhantes encontrados nos mesmos. Tal estratégia é útil para tratar cada grupo de forma específica pelos algoritmos de mineração de dados. A seguir, são descritos alguns trabalhos que exemplificam esses objetivos.

Em (PANAPAKIDIS; PAPAGIANNIS, 2014) é proposta uma metodologia para predição de curto prazo que utiliza clusterização para agrupar perfis de consumo diário semelhantes. Nesse estudo, dois diferentes modelos foram testados, chamados de modelo A e modelo B. O primeiro utiliza uma rede neural artificial com o algoritmo *backpropagation* resiliente e agrupa o consumo diário através de uma codificação binária, 1- para feriados e 0- para dias normais. Já o modelo B possui as mesmas entradas do modelo A, com exceção de que o valor binário é substituído por um *label* que identifica o perfil de consumo do feriado. No modelo, cada grupo de feriados é submetido a uma rede neural diferente. Além dos valores de consumo e da identificação do perfil de consumo diário, a temperatura diária e algumas variantes são utilizadas como entradas dos algoritmos de predição dos dois modelos. Comparando os dois modelos, o modelo B demonstrou maior robustez nos resultados, o que reforça a melhoria da eficácia dos resultados quando se utiliza informações sobre os dias especiais (feriados).

Um modelo de predição de consumo de médio e longo prazo é proposto em (CHUNSHAN; XIAOFENG, 2015). Nele, indicadores socioeconômicos são submetidos a um algoritmo de clusterização que verifica a distância que cada indicador tem em relação ao consumo de energia, fazendo com que apenas os indicadores mais próximos da variável de

consumo sejam selecionados como entradas para o algoritmo de previsão. Para análise de clusterização, vinte e dois indicadores foram testados utilizando três diferentes distâncias, que foram: chebychev, euclidiana e Manhattan. Como resultado dessa fase, os indicadores população total, população rural e índice de preço tiveram menor relevância sobre o consumo. Já os indicadores PIB, exportações e importações tiveram maior influência sobre o consumo. A partir daí, essas variáveis foram submetidas a uma rede neural auto regressiva para realizar a previsão de longo prazo e os resultados foram considerados satisfatórios para concessionária de energia estudada.

Em (PANAPAKIDIS et al., 2015), o modelo de clusterização minCentropia foi aplicado para treinamento de um conjunto de dados visando melhorar o resultado de uma predição de consumo, onde subconjuntos de dados foram gerados. As entradas para o algoritmo de clusterização foram: valores de consumo diário em diferentes defasagens, temperaturas mínima e máxima diária, identificação do dia da semana e se o dia é feriado ou não, e o valor do consumo das 24h do dia a ser previsto. Para cada subconjunto gerado, representados por centróides, uma rede neural diferente é implementada e valores de consumo são preditos. Para comparar os resultados obtidos com esse modelo, chamado de modelo B, um modelo A foi desenvolvido sem a etapa de clusterização e, conseqüentemente, apenas uma RNA foi implementada. Em todos os testes realizados, o modelo B gerou resultados mais satisfatórios, o que evidencia mais uma vez a eficácia da utilização de clusterização em estratégias de predição de consumo.

Em (QUILUMBA et al., 2015) é proposta uma estratégia de predição de consumo de curto prazo que tem em sua etapa de pré-processamento a clusterização do consumo de clientes com perfis de consumo diário semelhantes. O algoritmo de predição é aplicado para cada grupo gerado e o resultado de cada um é somado a fim de obter a previsão global do sistema. A técnica de clusterização adotada foi a *k-means* e o algoritmo de predição utilizado foi uma rede neural, cujas entradas foram: o consumo de energia por consumidor, informações do calendário, como o dia e o mês do consumo e se o dia é feriado ou não, e a temperatura. Esse trabalho demonstrou que a utilização de informações por cliente podem ser uma alternativa satisfatória para previsão de consumo global do sistema.

O *Mean Shift* é uma técnica de clusterização não-paramétrica (FUKUNAGA; HOSTETLER, 1975), que diferente do *K-Means* (QUILUMBA et al., 2015), não requer conhecimento a priori do número de *clusters* e não restringe a forma dos mesmos. Através da utilização do *Mean Shift*, é possível estimar a função densidade de probabilidade (fdp) de um

*cluster*, através de propriedades como a moda ou a linha principal do mesmo. Durante esse processo, novos pontos (chamados de virtuais) pertencentes a um *cluster* são gerados, estratégia essa chamada de densificação, que foi utilizada, primeiramente, por (MIRANDA et al., 2012) para preencher com valores virtuais conjuntos de dados que são considerados escassos. É importante ressaltar que, no domínio de sistemas elétricos, a escassez dos dados é resultado de diversos fatores, como: má geração dos dados, presença de anomalias nos dados que impactam na modelagem dos sistemas e pela ausência de um histórico de dados adequados para o estudo do modelo.

O *Mean Shift* é amplamente utilizado na literatura em aplicações relacionadas com processamento e segmentação de imagens, como pode ser visto em (SINDHUJA, 2015), (MURE et al., 2016) e (KILIC et al., 2016). Já sua aplicação no setor elétrico possui poucas referências na literatura, como em (MIRANDA et al., 2012), onde é proposto um modelo para diagnóstico de falhas incipientes, que é baseado na análise de gases que se encontram dissolvidos em transformadores. Nesse modelo, RNAs autoassociativas são treinadas para que se tornem ajustadas para uma determinada condição de falha dos transformadores. A técnica *Mean Shift* é aplicada nesse modelo para realizar a densificação dos dados, já que a base de dados existente é considerada escassa e uma maior quantidade de informação é necessária para treinamento das RNAs, reservando os dados reais para a validação do processo. A cada entrada de dados, as RNAs geradas disputarão entre si para verificar qual realiza o reconhecimento mais próximo de um diagnóstico de falha.

Outra referência na área do setor elétrico encontrada é em (WANG; ZHENG, 2011), onde é definida uma estratégia utilizada em aplicações de monitoramento de consumo, na qual se deseja identificar aparelhos elétricos em funcionamento em ambientes residenciais. As informações sobre operações realizadas pelos aparelhos ficam armazenadas em um medidor. Como essas operações se repetem algumas vezes, sua presença se torna muito constante no sistema, o algoritmo do *Mean Shift* é utilizado para clusterizar essas informações visando o aumento da acurácia dessa identificação.

Em (SUMAILI, 2011), o *Mean Shift* é utilizado para a classificação de padrões de consumo de clientes de uma concessionária de energia, estratégia bastante utilizada para melhor definição de tarifas. Além da busca pelos padrões existentes no consumo, o *Mean Shift* também possibilitou a descoberta de *outliers* na base de dados analisada. Para fins de comparação do desempenho do *Mean Shift*, foi aplicada também a técnica *K-Means* para realizar essa clusterização.

A partir da análise dos artigos apresentados nesta seção, nota-se que as técnicas de clusterização são amplamente empregadas em estratégias de predição de consumo e um dos modelos mais utilizados é o *K-Means*. Porém, a aplicação do K-Means, bem como de outros modelos encontrados na literatura, necessita que o número de *clusters* seja definido antes da execução do algoritmo, o que representa uma desvantagem, já que pode limitar a aplicação desses modelos apenas a bases de dados consideradas “comportadas”, ou seja, com pouca presença de *outliers*.

No estudo de caso trabalhado nesta tese, em virtude de haver uma grande presença de valores de consumo em dias especiais (nesse caso, feriados), a utilização de algoritmos de clusterização pode ser um fator relevante para descoberta do padrão de comportamento do consumo nesses dias. Adicionalmente a isso, o histórico de consumo nesses dias é limitado, ou seja, esses eventos estão em pequena quantidade na base de dados, dificultando ainda mais o processo de aprendizagem dos algoritmos, aumentando assim a dependência que o modelo tem do especialista. Devido a esses fatores, a utilização do algoritmo ITL *Mean Shift* para a clusterização, que possui a identificação da quantidade de *clusters* independente do analista, representa um dos diferenciais desse trabalho.

Além desse, aponta-se como outro diferencial, a própria aplicação do ITL *Mean Shift* em modelos para predição de consumo, que conforme visto nesta seção, tem sido utilizado em estratégias de clusterização e densificação mas em outros domínio de aplicação.

A Tabela 3.1 apresenta de forma resumida os principais estudos apresentados nesse capítulo, bem como o diferencial que o modelo proposto neste trabalho tem sobre eles.

Tabela 3.1 – Resumo dos trabalhos correlatos

Item	Autores das Referências Pesquisadas	Referências Pesquisadas	Diferencial da tese
1	(SHAO et al., 2015), (GARULLI et al., 2015), (SAIPUNYA et al., 2014), (PAPADOPOULOS; KARAKATSANIS, 2015).	Predição de consumo de curto prazo utilizando informações exógenas ao sistema elétrico.	Utilizar apenas informações relacionadas com o consumo.
2	(RAZA et al., 2014), (LAHOUAR; SLAMA, 2015), (SAIPUNYA et al., 2014), (ARORA; TAYLOR, 2013).	Formas subjetivas ou exatas para descoberta dos padrões de consumo dos dias especiais.	Algoritmos para descobertas automáticas dos padrões de consumo dos dias especiais.
3	(SINDHUJA, 2015), (MURE et al., 2016), (KILIC et al., 2016), (MIRANDA et al., 2012), (WANG; ZHENG, 2011), (SUMAILI, 2011).	Aplicação do ITL MS em processamento de imagens, monitoramento de carga e tarifação do setor elétrico.	Aplicação do ITL MS em estratégias de predição de carga.

4	(PANAPAKIDIS; PAPAGIANNIS, 2014), (PANAPAKIDIS et al., 2015), (WANG, 2015), (CHUNSHAN; XIAOFENG, 2015), (CUGLIARI et al., 2016).	Aplicação de técnicas de clusterização (K-Means e SOM) em estratégias de predição de carga.	Utilização do ITL MS para clusterização.
---	--	---	--

### 3.4 CONSIDERAÇÕES FINAIS

Neste capítulo, foram apresentados trabalhos que ilustram o estado-da-arte de técnicas e modelos empregados em predição de energia elétrica, a fim de demonstrar a diversidade de estratégias que vem sendo aplicadas para esse fim. Objetivou-se também demonstrar possíveis lacunas existentes na literatura que poderão ser utilizadas como objeto de estudo desta tese.

A partir das análises descritas, é possível verificar que os estudos propostos nesta tese, cujo objetivo é construir um modelo de predição de consumo de curto prazo que utiliza a técnica ITL *Mean Shift* para estudo e tratamento dos dias especiais, e algoritmos de predição, como RNAs e RLM, visando a melhorar a acurácia da predição, representam uma demanda importante a ser aplicada no domínio de sistemas de energia.

Uma consideração importante é que um dos diferenciais desse trabalho é a utilização do algoritmo ITL Mean Shift para descoberta dos padrões de consumo nos feriados através da clusterização dos dados. O ITL MS foi escolhido como método de clusterização pois os mesmo não precisa do conhecimento a priori dos clusters existentes e também pela tarefa de densificação dos dados, que ocorre ao mesmo tempo em que clusteriza os dados. Essas características não são realizadas pelos algoritmos de clusterização encontrados na literatura, como Kohonen e K-Means.

## **4. METODOLOGIA DE PREDIÇÃO DE CURTO PRAZO COM TRATAMENTO PARA OS DIAS ESPECIAIS**

### **4.1 CONSIDERAÇÕES INICIAIS**

A proposição de modelos para previsão de consumo de energia elétrica está em constante evolução e vem se consolidando como uma área de intensa investigação pela comunidade científica, devido, entre outros aspectos, aos interesses operacionais, legais e financeiros, os quais motivam as concessionárias a aperfeiçoar continuamente o desempenho de suas estratégias de previsão. Para tanto, uma grande variedade de técnicas tem sido desenvolvida e empregada com relativas vantagens e desvantagens, de acordo com cada tipo de aplicação.

Diante da importância que os modelos de previsão tem para o setor elétrico, esta tese propõe o desenvolvimento de uma nova metodologia para previsão de consumo de energia, inovadora e factível, que inclui em seu algoritmo de previsão informações acerca dos dias especiais, bem como densifica esses valores na base de dados de consumo, com o intuito de melhorar o aprendizado dos relacionamentos existentes nos dados. Os detalhes de seu funcionamento encontram-se descritos nas seções a seguir.

Este capítulo está dividido em mais duas seções, a saber: a primeira, que faz uma introdução sobre o contexto da problemática estudada; e a segunda, que apresenta com detalhes a metodologia que soluciona a problemática.

### **4.2. INTRODUÇÃO DA PROBLEMÁTICA**

Os estudos demandados neste trabalho se originaram a partir das necessidades reais da concessionária de energia elétrica do Estado do Pará – CELPA de melhorar sua capacidade de planejar a compra de energia para o mercado futuro. Os métodos utilizados pela concessionária para previsão de curto prazo geravam erros em torno de 4%, valor considerado alto do ponto de vista financeiro pela companhia.

A partir desse contexto, com a execução do projeto de pesquisa “ANEEL/CELPA Predict - Ferramenta de Suporte à Decisão para Predição de Cargas de Sistemas Elétricos”, métodos matemáticos e de inteligência computacional foram implementados para prever as

necessidades de compra de energia no mercado futuro e para realizar inferências sobre a situação do sistema elétrico da CELPA, a partir de dados históricos de consumo e suas correlações com dados socioeconômicos e climáticos.

No decorrer da execução do Predict foi constatado, de maneira inequívoca, o efeito que o consumo em dias especiais imprimiam nas previsões e correlações realizadas. Dessa forma, constatou-se a necessidade de aprender sobre a influência que esses dias têm sobre o consumo e se a inclusão desse conhecimento em um modelo de predição contribui para melhoria dos resultados de previsão da concessionária. Essa necessidade, portanto, motivou o desenvolvimento da metodologia proposta nesse trabalho, que visa realizar predição de consumo de curto prazo (previsão para o próximo dia) utilizando o conhecimento sobre os dias especiais (feriados, especificamente) adquiridos através da clusterização. Destaca-se que o treinamento do algoritmos de previsão utiliza os valores virtuais de consumo dos feriados, obtidos através da densificação.

Por intermédio da utilização de algoritmos de clusterização, pretende-se adquirir conhecimentos acerca do comportamento dos feriados. Estratégia essa considerada importante para o aprimoramento dos resultados nesse horizonte de predição pois a ocorrência desses dias impacta diretamente no consumo de dias próximos aos mesmos. A densificação dos valores de consumo diário dos feriados, na qual novas amostras são geradas (valores virtuais), é um importante artifício para melhorar o treinamento dos algoritmos de predição propostos e consequentemente gerar melhores resultados, já que com uma quantidade maior de amostras, esses algoritmos são capazes de aprender melhor o relacionamento existente entre os dados de consumo. Um outro aspecto importante sobre a densificação de valores insuficientes nas bases de dados é a possibilidade de aplicá-los também para o estudo de outros tipos de padrões existentes, também inerentes ao sistema elétrico, como os apagões (GALL, 2002).

O conhecimento adquirido sobre os dias especiais por meio da aplicação da metodologia proposta pode ser estendido para outros horizontes de previsão e também para outros propósitos no setor elétrico, como operação, manutenção, supervisão e novos serviços em ambientes *smart grids* (e.g. *smart home*), que assim como ocorre no planejamento, também são muito suscetíveis à variações ocorridas na curva diária de consumo.

#### 4.2.1. PREDIÇÃO DE CONSUMO

Previsão de consumo é um elemento chave em processos de tomada de decisão em sistemas elétricos, uma vez que qualquer decisão depende fortemente dos eventos que antecedem o consumo. Desta forma, é de grande interesse o estudo de métodos de previsão que consigam identificar as estruturas e padrões existentes nos dados históricos e que sejam capazes de gerar os próximos padrões da série. No entanto, em muitos problemas reais, a ocorrência de bases de dados com informações insuficientes acerca de determinados fenômenos impedem que os algoritmos de previsão tenham resultados satisfatórios.

Desta forma, a busca por modelos de previsão capazes de tratar dados insuficientes (consumos em feriados, por exemplo) e de aprender a influência que essas informações têm sobre o consumo, favorecendo assim melhores resultados de previsão, pode prover um grande diferencial financeiro para as concessionárias de energia elétrica.

#### 4.2.2. DIAS ESPECIAIS E GERAÇÃO DE VALORES VIRTUAIS

As previsões de curto prazo sofrem muita influência de diferentes padrões de consumo, de acordo com tipo de dia (e.g dias úteis, fins de semana e feriados) (RAZA et al., 2014). Neste contexto, o estudo do consumo nos feriados tem sido uma estratégia de destaque na literatura especializada, conforme descrito no Capítulo 3. Sua importância para os modelos de previsão advém do fato da sua ocorrência alterar os padrões de consumo dos dias adjacentes a ele, situação essa bastante intensificada dependendo do dia da semana da ocorrência desse feriado. Diante da importância da análise dos feriados, na metodologia proposta os dias especiais a serem analisados serão os feriados.

Um ponto importante a ser considerado no estudo dos feriados é a dificuldade encontrada para sua análise devido a seu número reduzido de ocorrências - cada feriado ocorre uma vez por ano e o número de anos históricos que podem ser efetivamente usados para a previsão de consumo é limitada. Além dessa limitação, outra consideração importante sobre a presença dos feriados, é que a maioria deles podem ocorrer em diferentes dias da semana (são poucos os que sempre ocorrem no mesmo dia da semana, como o feriado da sexta-feira santa, por exemplo). Esse fato restringe ainda mais a sua análise pois, conforme será visto a seguir, dependendo do dia da semana em que ocorrem, o feriado influencia de forma diferente no consumo de seus dias adjacentes.



Com base no exposto, é importante que valores virtuais de consumo em feriados sejam gerados para que os algoritmos de predição possam aprender seu padrão de comportamento na série de consumo.

### **4.3 DESCRIÇÃO DA METODOLOGIA PROPOSTA**

Este trabalho, visando a contribuir com soluções para a problemática levantada nas seções anteriores a este capítulo, propõe uma metodologia de predição inovadora e factível, que inclui em seu algoritmo de previsão informações acerca dos feriados, bem como densifica esses valores na base de dados de consumo, com o intuito de melhorar o aprendizado dos relacionamentos existentes nos dados.

Nas análises propostas neste trabalho, serão utilizados dados provenientes da base de dados de consumo de energia da CELPA compreendidas entre o período de janeiro de 2002 a setembro de 2012. Nas análises seguintes, os dados de consumo diário serão representados por  $\{X_t | 1 \leq t \leq m\}$ , onde  $X$  refere-se ao consumo do dia  $t$  e  $m$  é o número de amostras de consumo diário da base de dados. O fluxograma da metodologia proposta é demonstrado na Figura 4.1.

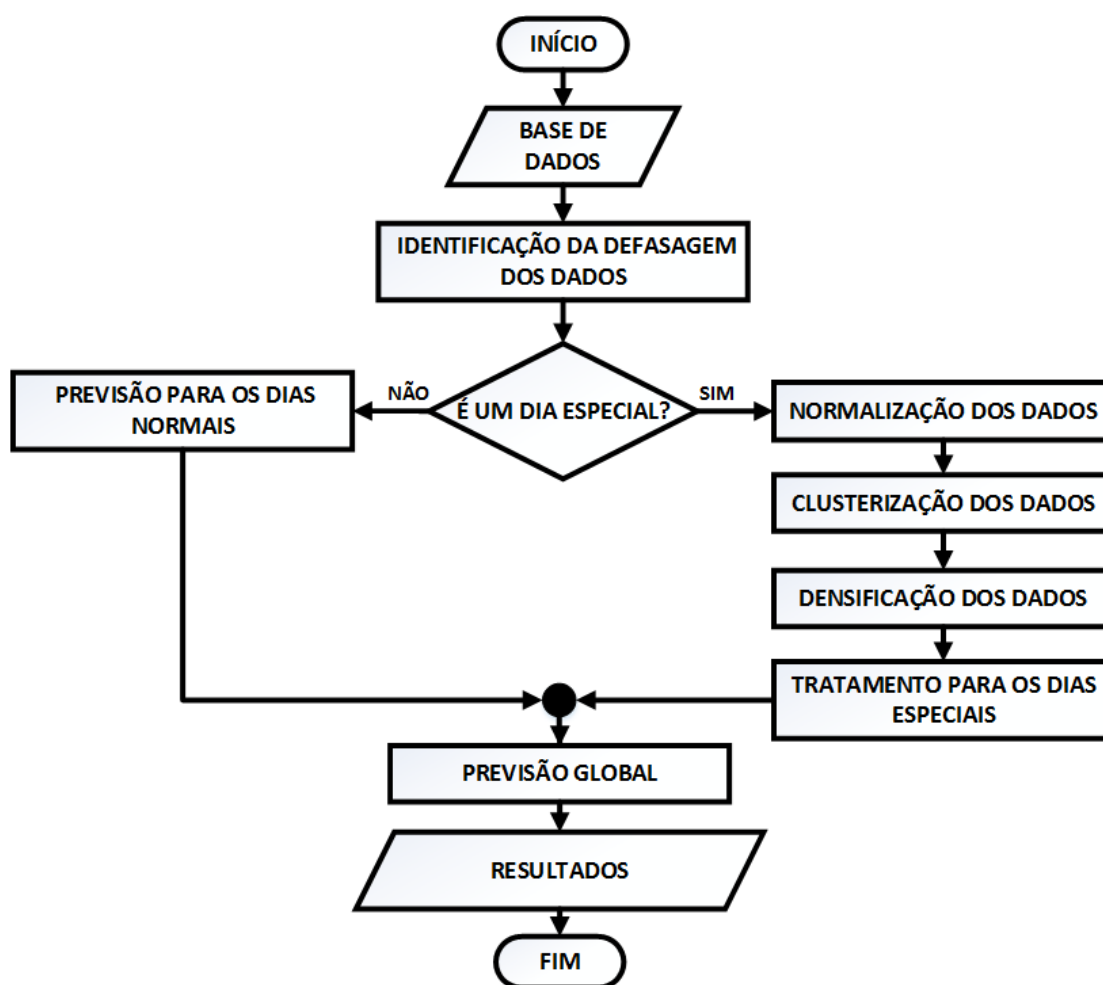


Figura 4.1 – Metodologia de Predição de Consumo.

No fluxograma da Figura 4.1, o bloco “Base de dados” refere-se às entradas dos algoritmos implementados na metodologia. A Seção 4.3.1 descreve a etapa de “Identificação da defasagem”. Continuando o fluxo, existe a necessidade de uma decisão que visa a verificar se o dia a ser predito é feriado ou não, dado que esses dias são tratados de forma diferente pela metodologia. Caso seja um dia normal, os algoritmos da etapa de “Predição para os dias normais” (Seção 4.3.6) serão executados, caso contrário, o fluxo seguirá para as etapas de “Normalização dos dados”, “Clusterização dos dados”, “Densificação dos dados” e “Tratamento para os dias especiais”, discutidas nas Seções 4.3.2, 4.3.3, 4.3.4 e 4.3.5, respectivamente. E, finalmente, o fluxo segue para a etapa de “Predição Global” (Seção 4.3.7), cujos resultados estão representados pela etapa “Resultados”. A seguir, cada uma dessas etapas serão detalhadas.

#### 4.3.1 IDENTIFICAÇÃO DA DEFASAGEM DOS DADOS

Uma vez que a previsão do consumo será feita baseada principalmente em medições de consumo anteriores, é preciso então definir quantos dias são necessários para obter os melhores resultados de predição, ou seja, é necessário descobrir qual o número de defasagens que melhor explica a variável dependente (nesse caso, o consumo de uma feriado). Com esse propósito, a métrica do coeficiente de determinação ( $R^2$  – correlação múltipla quadrada) (SANTANA, 2008).

A métrica  $R^2$  afere o poder explicativo de uma regressão (cujo modelo foi descrito na Seção 2.4.3), ou seja, a capacidade das variáveis independentes ( $X_i$ ) em explicarem as variações na variável dependente  $Y_i$ . O  $R^2$ , também chamado de coeficiente de determinação múltipla, é calculado através da equação a seguir:

$$R^2 = \frac{A^t(X^tX)A - n\bar{Y}^2}{y^ty} \quad (4.1)$$

onde  $\bar{Y}$  é o valor da variável  $Y$ , e  $y$  a variável reduzida e centrada na média, obtida pela subtração da média  $\bar{Y}$  de cada observação  $Y_i$ ,  $y = (Y_i - \bar{Y})$ .

A métrica  $R^2$  foi adotada para avaliar quantos  $p$  (tamanho da janela) dias anteriores ao feriado em questão melhor explica o consumo nesse feriado. Então, se  $p$  é o tamanho da janela, para estimar  $X_{t+1}$  os seguintes valores de consumo serão necessários:  $X_{t-p+1}, X_{t-p+2}, \dots, X_{t-1}, X_t$ . Neste trabalho, chamar-se-á de série de consumo defasado  $S_i$  o conjunto de dados formado por  $X_{t+1}$  e seus valores defasados, isto é,  $(X_{t-p+1}, X_{t-p+2}, \dots, X_{t-1}, X_t)$ , ficando  $S_i = X_{t-p+1}, X_{t-p+2}, \dots, X_{t-1}, X_t, X_{t+1}$ , com  $\{S_i | 1 \leq i \leq m\}$ . O resultado obtido será empregado nos estágios seguintes para prever os valores de consumo para o próximo dia, que poderá ser um feriado ou não.

#### 4.3.2 NORMALIZAÇÃO DOS DADOS

A normalização visa a transladar os valores de uma base de dados para um mesmo nível de grandeza, mantendo-os no mesmo intervalo. Nesta etapa, cada série de consumo defasado  $S_i$  será normalizada, visando contribuir para melhorar a acurácia da etapa de clusterização.

Sendo assim, cada valor de consumo  $X_a | t - p + 1 \leq a \leq t + 1$  presente em  $S_i$  é dividida pelo valor máximo do consumo em  $S_i$ , conforme mostra a Equação 4.1.

$$X_a^N = \frac{X_a}{\max(S_i)} \quad \forall a | t - p + 1 \leq a \leq t + 1 \quad (4.1)$$

Onde  $p$  é o tamanho da janela (conforme seção 4.3.1),  $X_{t-p+1}$  é o consumo do  $n$ -ésimo dia antes do dia  $t+1$ , e assim sucessivamente. A série de consumo defasado normalizada é representada por  $S_i^N$  e seus valores ficam dentro do intervalo  $[0,1]$ .

Para esta tese, a importância de normalizar as variáveis de entrada para a clusterização é fazer com que os valores de consumo estejam em um mesmo nível de grandeza a fim de possibilitar a análise de similaridades entre os dados e a busca por padrões existentes entre os mesmos. De outra forma, iria enviesar os resultados gerando *clusters* incorretos.

### 4.3.3 CLUSTERIZAÇÃO DOS DADOS

O objetivo desta etapa é descobrir os padrões existentes nos valores de consumo diários presentes na base de dados estudada. Mais especificamente, pretende-se identificar como os feriados se relacionam entre si, bem como, o relacionamento de tais feriados com os dias normais.

A identificação do *cluster* ao qual pertence determinado consumo será utilizada como variável de entrada dos algoritmos de predição, devido à influência que o dia da ocorrência do feriado tem sobre o consumo. A informação sobre qual *cluster* o feriado pertence vai auxiliar esses algoritmos na melhoria do aprendizado do comportamento da série temporal.

Para essa análise de similaridade, a técnica utilizada será o ITL MS, onde as entradas serão conjuntos  $S_i^N$ , em que  $X_{t+1}^N$  representa o consumo de um feriado. A saída esperada são os conjuntos  $S_i^N$  agrupados conforme um padrão de consumo semelhante.

Os resultados gerados pelos algoritmos de clusterização devem ser avaliados para que os *clusters* gerados sejam validados ou não, ou seja, é importante verificar se os conjuntos foram agrupados de forma correta. Conjuntos agrupados de forma equivocada pelo algoritmo podem necessitar de ajustes feitos pelo analista para identificação correta do *cluster* a que pertencem esses conjuntos. Ressalta-se que nos algoritmos de predição só serão utilizadas informações referentes a *clusters* validados.

### 4.3.4 DENSIFICAÇÃO DOS DADOS

O processo de densificação é um dos resultados obtidos pelo ITL MS, que gera valores (virtuais) que possuem as mesmas propriedades estatísticas das amostras reais presentes nas bases de dados. Neste trabalho, a densificação será utilizada para gerar valores virtuais do

consumo em feriados, visando melhorar o treinamento dos algoritmos de predição de consumo e, conseqüentemente, obter erros menores de predição.

Como foi explicado na Seção 2.3.2, a densificação ocorre enquanto o processo de clusterização é realizado pelo ITL MS, onde novas amostras são geradas para cada *cluster*. Então, para cada *cluster* tem-se  $A = \{S_1^N, S_2^N, \dots, S_k^N\}$ , que agrupa um determinado  $k$  número de séries de consumo defasado, um conjunto  $A_v = \{SV_1^N, SV_2^N, \dots, SV_d^N\}$  será criado com  $d$  séries virtuais defasadas ( $SV_i$ ) com características similares àquelas do grupo  $A$ . Cada *cluster* densificado será então  $\{A \cup A_v\}$ .

#### 4.3.5 TRATAMENTO PARA OS DIAS ESPECIAIS

A pequena quantidade de amostras de consumo nos feriados em bases de dados de consumo é um fator que dificulta o aprendizado de seu comportamento, bem como também o aprendizado sobre a influência que esses dias têm sobre os dias que o cercam.

Nesta etapa, baseado no conhecimento com a clusterização e com os valores virtuais obtidos com a densificação, os algoritmos de predição são aplicados para construir modelos de estimação para cada padrão de feriado (para cada *cluster*). A entrada para os algoritmos são os valores de consumo da série  $S_i^N$  que foi usada no ITL MS, agora fixando  $X_{t+1}^N$  como variável de saída.

Objetivando melhorar os resultados gerados pelo algoritmo de predição, na fase de treino serão empregados apenas valores virtuais de consumo, reservando os valores reais para a fase de teste.

É importante destacar que os valores virtuais de consumo gerados pela densificação não podem ser utilizados da forma como foram gerados pelos algoritmos de predição, uma vez que com o processo de normalização a tendência de crescimento da série de dados foi perdida. Portanto, para possibilitar a aplicação desses valores em estratégias de previsão de consumo é preciso utilizar alguma estratégia que devolva a tendência de crescimento da série aos dados, que é o que acontece na etapa de Predição Global.

#### 4.3.6 PREVISÃO PARA OS DIAS NORMAIS

Esta etapa consiste no estabelecimento de parâmetros que define o modelo de predição do consumo somente para os dias normais. Neste trabalho serão empregados os algoritmos RNA e RLM para previsão, dada sua ampla utilização nesse tipo de problema.

Os algoritmos de predição usarão a série de consumo  $S_i^N = X_{t-p+1}^N, X_{t-p+2}^N, \dots, X_{t-1}^N, X_t^N, X_{t+1}^N$ , onde  $t + 1$  é um dia normal. A entrada para os algoritmos é a série de dados organizada conforme o tamanho da janela  $p$ ,  $X_{t-p+1}^N, X_{t-p+2}^N, \dots, X_{t-1}^N, X_t^N$ , e a saída é o consumo predito para um dia normal ( $X_{t+1}^N$ ).

#### 4.3.7 PREDIÇÃO GLOBAL

O processo de predição global resulta da combinação dos modelos para os dias normais e feriados. O objetivo desta etapa é construir a série de consumo com os dois tipos de dias, bem como sua tendência de crescimento. Isso se deve à ausência de conhecimento temporal sobre o consumo gerado pela etapa de normalização dos dados.

Esta etapa usará os melhores modelos encontrados durante as análises com os feriados e dias normais (Seções 4.3.5 e 4.3.6, respectivamente). Nesta etapa, os modelos serão alternados conforme o tipo de dia a ser predito: quando se deseja prever o consumo de um feriado, o modelo para esses dias será empregado, caso contrário, o modelo para prever dias normais será utilizado. Para prever o consumo de mais de um dia à frente, o algoritmo de previsão utiliza como entrada somente um valor de consumo real. À medida que ele vai prevendo um valor de consumo, esse valor será utilizado para prever o consumo do próximo dia, como em um processo de realimentação, até a finalização do teste.

### 4.4 CONSIDERAÇÕES FINAIS

Neste capítulo, inicialmente, houve a contextualização da metodologia proposta no cenário de predição de consumo, bem como a sua importância na tentativa de obter melhores resultados de predição que os obtidos pelos métodos já utilizados pela concessionária de energia estudada. Além disso, houve a apresentação da metodologia, onde cada uma de suas etapas foram detalhadas e o seu modelo foi construído.

No capítulo a seguir, a metodologia será aplicada em um estudo de caso para que sua eficácia seja avaliada.

## **5. APLICAÇÃO DA METODOLOGIA AO ESTUDO DE CASO DA CELPA**

### **5.1 CONSIDERAÇÕES INICIAIS**

Neste capítulo, a metodologia proposta, já descrita no capítulo anterior, é aplicada ao caso da CELPA, visando a aperfeiçoar os estudos iniciados com o projeto Predict e melhorar a acurácia dos modelos de predição já implementados pela empresa através da aplicação do conhecimento adquirido com as etapas de clusterização e densificação da metodologia.

Nas seções seguintes, cada etapa da aplicação da metodologia aos dados da CELPA será descrita, bem como também será apresentada a avaliação dos resultados obtidos. Na Seção 5.2.6 são discutidos aspectos relacionados com desempenho e escalabilidade da metodologia.

### **5.2 APLICAÇÃO DA METODOLOGIA**

#### **5.2.1 ANÁLISE DA BASE DE DADOS DE CONSUMO DA CELPA**

Conforme citado anteriormente, os estudos propostos aqui foram realizados utilizando a base de dados de consumo (energia requerida total<sup>1</sup>) da CELPA, a qual possui valores históricos de consumo diário de Janeiro de 2002 a Setembro de 2012. A Figura 5.1 representa a série histórica de consumo diário da CELPA para esse período.

Nessa base de dados, de um total 3275 amostras de consumo diário, existem 94 amostras de consumo em feriados, que representa 2,39% do total. O pequeno percentual de amostras de consumo em feriados nessa base de dados ressalta a dificuldade de análise desses dias quando comparada a quantidade de amostras que representam os dias normais de consumo. Os algoritmos de predição ao analisarem uma série como essa podem entender que

---

<sup>1</sup> Energia requerida total inclui o consumo de todas as classes de consumo existentes.

os feriados tratam-se de ruídos, ignorando a influência desses dias sobre o consumo.

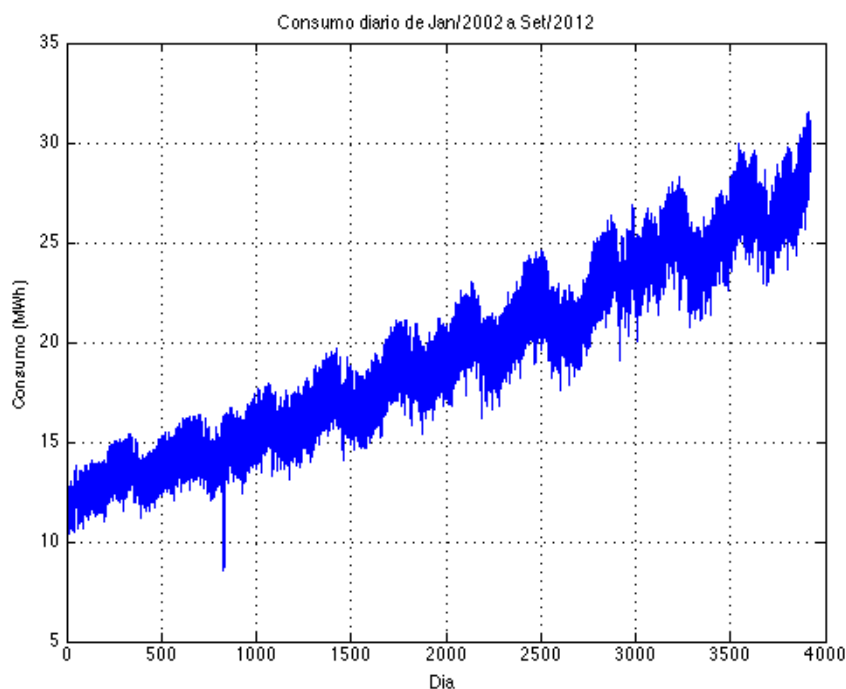
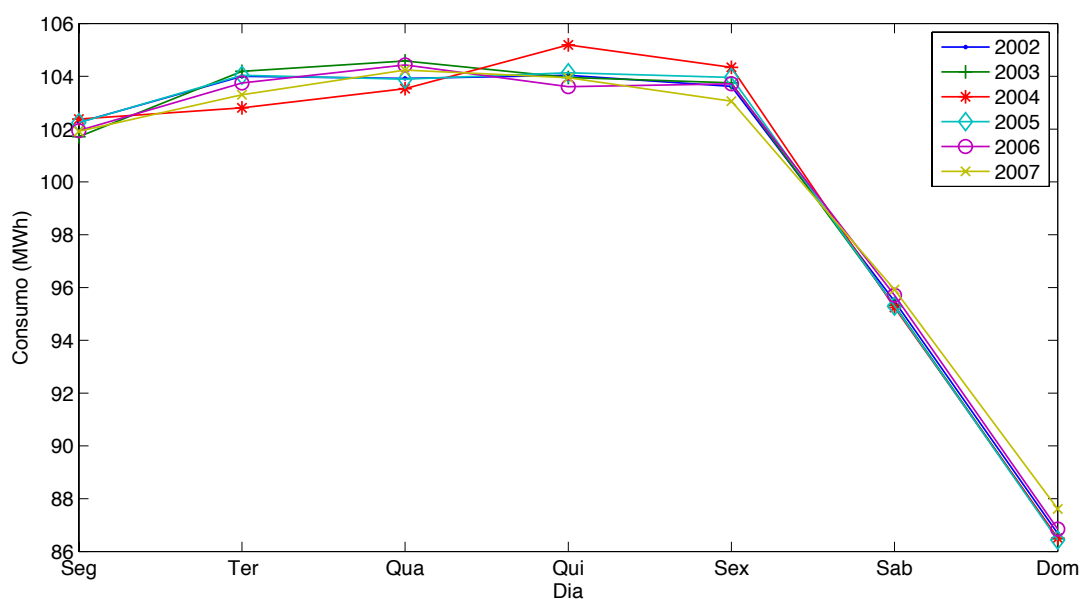


Figura 5.1 - Série histórica de consumo da CELPA.

A série histórica da CELPA, a partir da análise de seu correlograma, é caracterizada como explosiva e não estacionária. Essas características devem-se, principalmente a dois fatores: (1) a ocorrência do “apagão” no ano de 2002 e sua consequente campanha de racionamento de energia, que alterou o padrão da série para outros níveis de consumo; e (2) o crescimento econômico que o Estado do Pará apresentou nos últimos anos, que vem alterando os hábitos dos consumidores, e, conseqüentemente, os padrões de consumo da série. A característica do item 2 representa uma particularidade da série de consumo da CELPA, que a diferencia do padrão de consumo de séries de outras concessionárias do país. Ressalta-se que esses fatores dificultam ainda mais o aprendizado dos relacionamentos existentes nos dados e impactam sobremaneira na sua modelagem para predição.

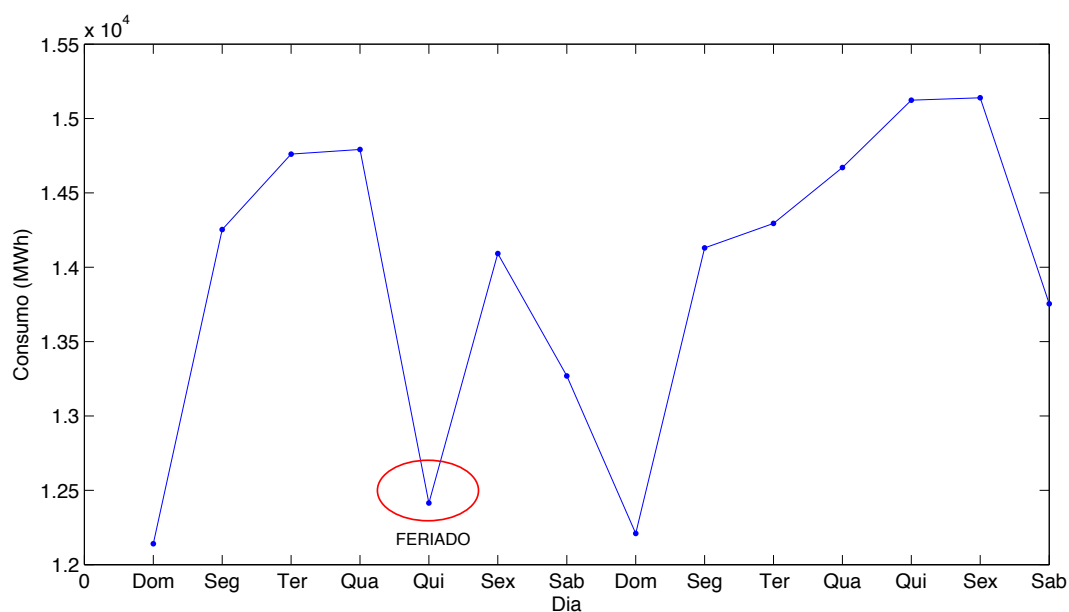
Visando entender o comportamento do consumo em uma semana sem feriado, o padrão de consumo semanal da série é mostrada na Figura 5.2. Nela, os pontos representam uma média dos valores de consumo normalizados pelo maior valor de consumo de série, para todas as semanas normais existentes (semanas sem feriados). A partir da análise dessa figura, é possível reconhecermos a similaridade existente nos padrões de consumo durante a semana (de segunda-feira a sexta-feira), sendo aproximadamente linear, e uma queda gradativa nos consumos de sábado e domingo.





**Figura 5.2** – Padrão de consumo semanal da série.

Portanto, é possível identificar o consumo de acordo com o dia da semana, onde os valores mais altos ocorrem durante a semana e os valores mais baixos nos finais de semana e feriados. A ocorrência de um feriado no final de semana não interfere no padrão de consumo de outros dias, mas quando um feriado ocorre no meio da semana, os dias adjacentes a ele tem seu consumo afetado. A Figura 5.3 mostra um exemplo desse comportamento, quando um feriado ocorre em uma quinta-feira.



**Figura 5.3** – Variações do consumo quando um feriado ocorre no meio da semana. Padrão de consumo semanal da série.

A série da Figura 5.3 representa o feriado do dia 1º de maio de 2003, que ocorreu em uma quinta-feira. Analisando a Figura 5.3 e comparando-a com a Figura 5.2, percebe-se que os consumos da sexta-feira e segunda-feira seguintes ao feriado possuem níveis de consumo bem abaixo do consumo realizados às segundas e sextas-feiras de uma semana sem feriados. Ou seja, o dia da semana em que o feriado ocorre reflete o comportamento não rotineiro dos consumidores, que muda seu padrão de consumo, e normalmente produz seus efeitos também nos dias adjacentes a ele.

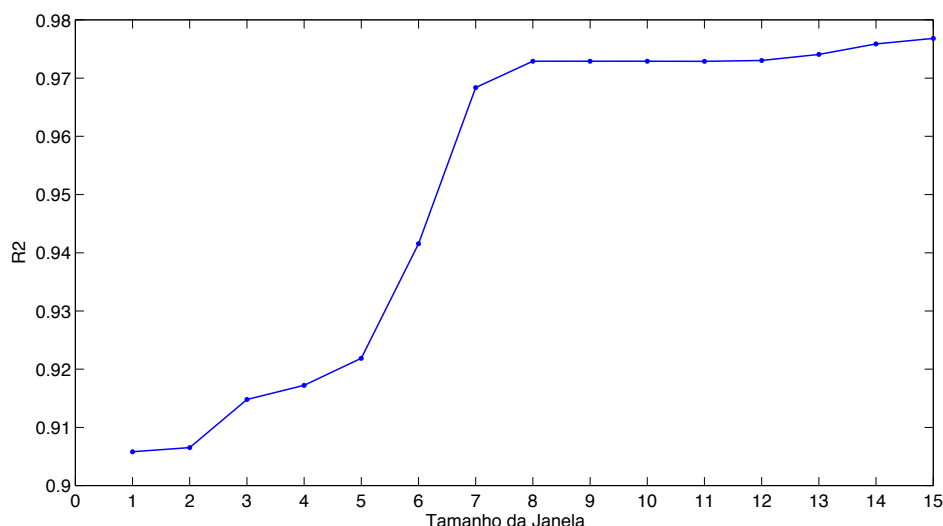
Apesar da ocorrência do feriado em determinada semana ser conhecida, é importante verificar se os algoritmos conseguem captar que essas variações de consumo são resultantes da ocorrência de um feriado. Neste trabalho, esse conhecimento é obtido através da aplicação das técnicas de clusterização, em que cada feriado estará ligado a um padrão em particular, ou *cluster*.

A clusterização dos feriados foi gerada através do algoritmo ITL MS e, para confirmar os resultados obtidos, a mesma análise de clusterização foi também implementada utilizando redes de *Kohonen*. Detalhes sobre o processo de clusterização, seus resultados e outros passos da metodologia serão explicadas nas seções seguintes.

## 5.2.2 IDENTIFICAÇÃO DA DEFASAGEM DOS DADOS

Nesta metodologia, é necessário definir qual o melhor conjunto de dados, composto por diferentes defasagens do consumo, serão usados para predição do consumo. Essa definição é realizada através da identificação do tamanho  $p$  da janela de defasagens.

Com esse propósito, vários experimentos foram realizados considerando defasagens de consumo de até 15 dias anteriores ao feriado (variável predita), usando a métrica do coeficiente de determinação  $R^2$ . Analisando os resultados, foi possível verificar que a janela de tamanho  $q=7$  é suficiente para explicar cerca de 97% da variável predita, conforme é visto na Figura 5.4. Essa figura também mostra que a partir do sétimo dia, o valor de  $R^2$  não melhora significativamente. Diante desse resultado, os padrões de entrada dos algoritmos serão então representados, para cada amostra, como uma sequência de 7 dias anteriores ( $X_{t-6}, X_{t-5}, \dots, X_t$ ) ao dia que será predito ( $X_{t+1}$ ).

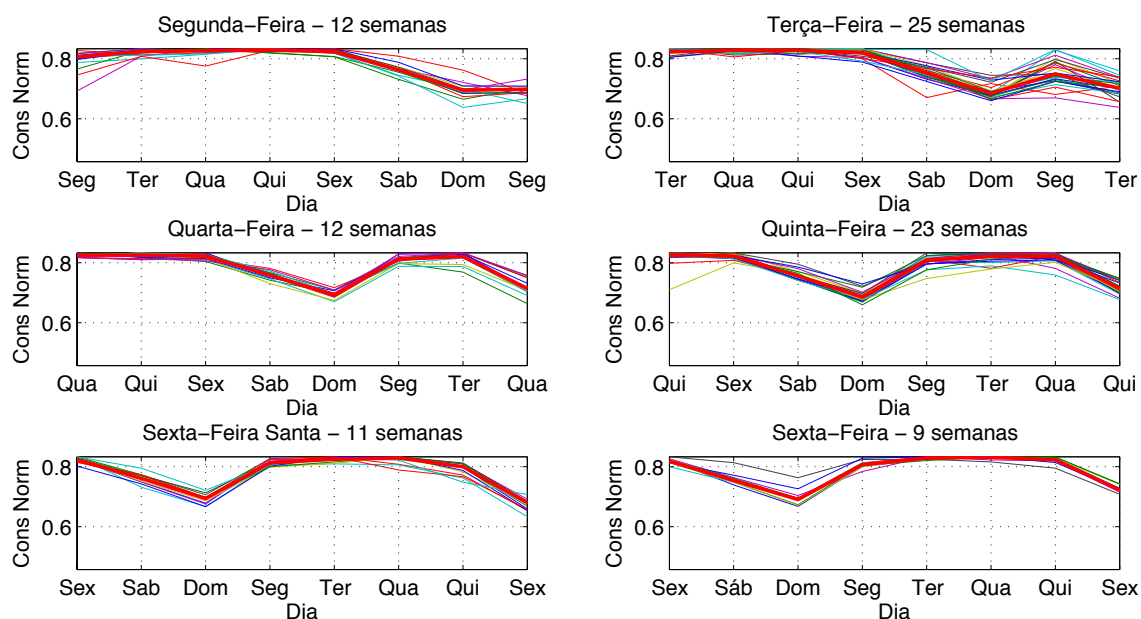


**Figura 5.4** – Coeficiente de determinação  $R^2$  para identificação da janela de defasagens.

### 5.2.3 CLUSTERIZAÇÃO DOS DADOS

Nesta etapa, apenas os conjuntos de dados com feriados são considerados, ficando as séries com o formato  $X_{t-p+1}^N, X_{t-p+2}^N, \dots, X_{t-1}^N, X_t^N, X_{t+1}^N$ , de acordo a defasagem com tamanho da janela  $p=7$ .

Utilizando as séries de consumo defasadas e normalizadas pelo valor máximo da janela de até 7 dias, o algoritmo ITL MS identificou 11 modas (*clusters*), com um valor de lambda  $\lambda = 0.5$  na Equação 2.15 da Seção 2.3.2. A Figura 5.5 apresenta os 6 principais *clusters* obtidos pela clusterização com ITL MS, correspondendo cada um ao padrão de um feriado ocorrido em um dos 5 dias da semana (de segunda-feira a sexta-feira – sendo que 2 *clusters* representam sextas-feiras). Os demais 5 *clusters* representam casos muito especiais de ocorrência de feriados e que não foram considerados para estudo dos dias especiais pois devem ser analisados individualmente.



**Figura 5.5** – Resultados da clusterização com ITL MS. 6 clusters foram descobertos, com 12 (segunda-feira), 25 (terça-feira), 12 (quarta-feira), 23 (quinta-feira), 11 (sexta-feira santa) e 9 (sexta-feira) semanas cada. As linhas coloridas representam o consumo real e a linha grossa vermelha representa a moda de cada cluster.

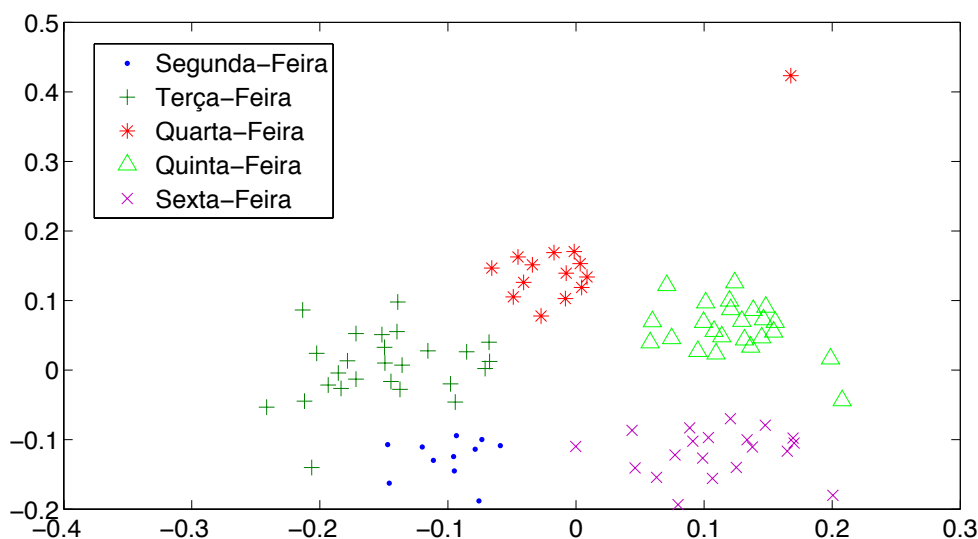
Assumindo que os valores mais baixos em cada *cluster* correspondem ao consumo em um domingo, é possível identificar que cada *cluster* está associado ao dia da semana em que o feriado ocorreu, sendo que o 8º valor da série corresponde ao consumo no feriado. É possível ver então que os 4 primeiros *clusters* representam o padrão com a ocorrência do dia especial de segunda-feira a quinta-feira, respectivamente. Além desses, o algoritmo ITL MS identificou 2 padrões diferentes para os feriados em sextas-feiras: o primeiro refere-se ao consumo nos feriados da sexta-feira santa (com 11 semanas) e o segundo representa os outros feriados desse dia (9 semanas). Essa separação ocorreu porque a quinta-feira que precede o feriado de sexta-feira santa possui uma queda significativa no consumo, queda essa explicada pelo não funcionamento de escolas e repartições públicas, que leva a um padrão ligeiramente diferente dos outros, uma vez que o mesmo não acontece com qualquer outro feriado que ocorre na sexta-feira.

No *cluster* de feriados na terça-feira, é possível verificar que o consumo na segunda-feira é, em média, menor do que nos outros dias da semana. Isto se deve ao fato de que na segunda-feira algumas lojas, repartições públicas e escolas não funcionam.

Conforme dito anteriormente, os 5 *clusters* remanescentes correspondem a casos especiais que devem ser analisados de forma mais criteriosa. Dos 5, 3 conjuntos referem-se ao feriado do ano novo, que apresenta um comportamento muito específico, dependendo do dia da semana em que ocorre; e 2 conjuntos representam o consumo no feriado em uma sexta-feira, que em virtude de uma queda anormal do consumo na segunda-feira e na terça-feira seguintes, o ITL MS os separou. Esses 2 últimos *clusters* podem representar eventos muito especiais que podem ser determinado pela concessionária em uma análise futura.

Para efeitos comparativos e também para ratificar os resultados obtidos pela clusterização com ITL MS, uma outra análise de clusterização foi implementada, agora utilizando uma rede de *Kohonen*, apresentado no Capítulo 2. Nessa análise, o mesmo conjunto de entrada do ITL MS foi utilizado.

Antes de aplicar rede de *Kohonen* na base de dados da CELPA, a técnica Mapas de Sammon (SAMMON, 1969) foi implementada com o objetivo de melhor visualizar a distribuição dos dados e definir os parâmetros do mapa de *Kohonen* mais claramente. Os resultados do mapa de Sammon podem ser vistos na Figura 5.6, onde se destaca a similaridade dos perfis de consumo de segunda-feira a sexta-feira, o que sugere a presença de 5 *clusters* nos resultados gerados pela rede de *Kohonen*.

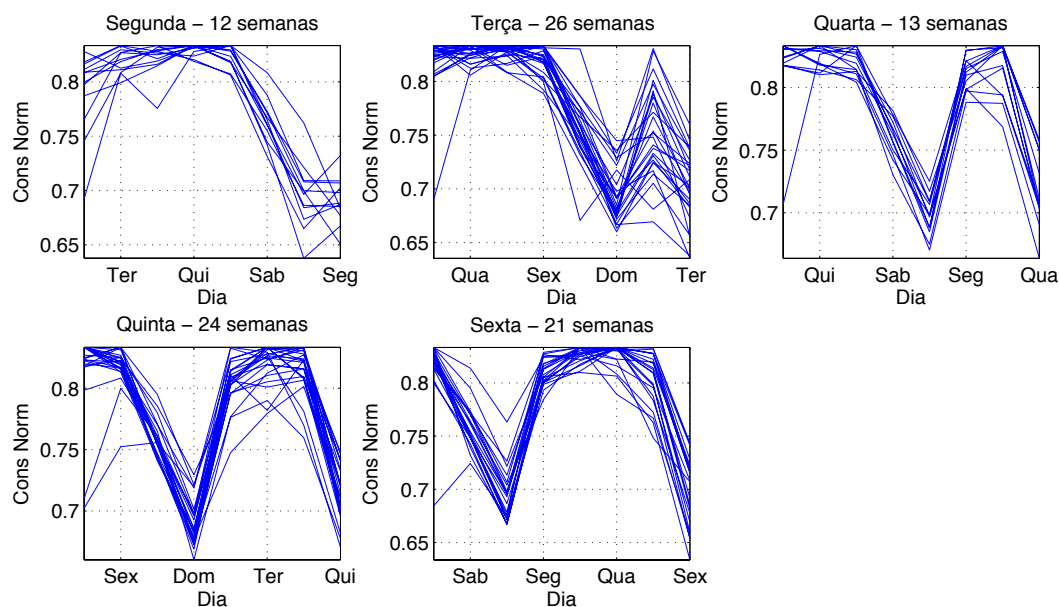


**Figura 5.6** – Resultados do mapa de Sammon

Para reforçar os resultados obtidos pelo mapa de Sammon, vários testes foram realizados para encontrar a melhor configuração (tamanho e topologia) da rede de *Kohonen*, onde o melhor resultado foi um mapa hexagonal 2x3. Nos testes com topologias de rede maiores (25x25, 10x10, 6x8 e 6x6, por exemplo) um grande número de neurônios foram

gerados para mapear a complexidade dos dados, resultando em *clusters* muito pequenos que representam um sobre ajustamento da rede neural.

Como aconteceu com os resultados gerados pelo ITL MS, a rede de *Kohonen* também separou padrões conforme o dia da semana em que o feriado ocorre, gerando assim 5 *clusters* principais, um para cada dia da semana, conforme visto na Figura 5.7.



**Figura 5.7** – Resultados do mapa Kohonen. 5 *clusters* foram descobertos, com 12 (segunda-feira), 26 (terça-feira), 13 (quarta-feira), 24 (quinta-feira) e 21 (sexta-feira) semanas cada.

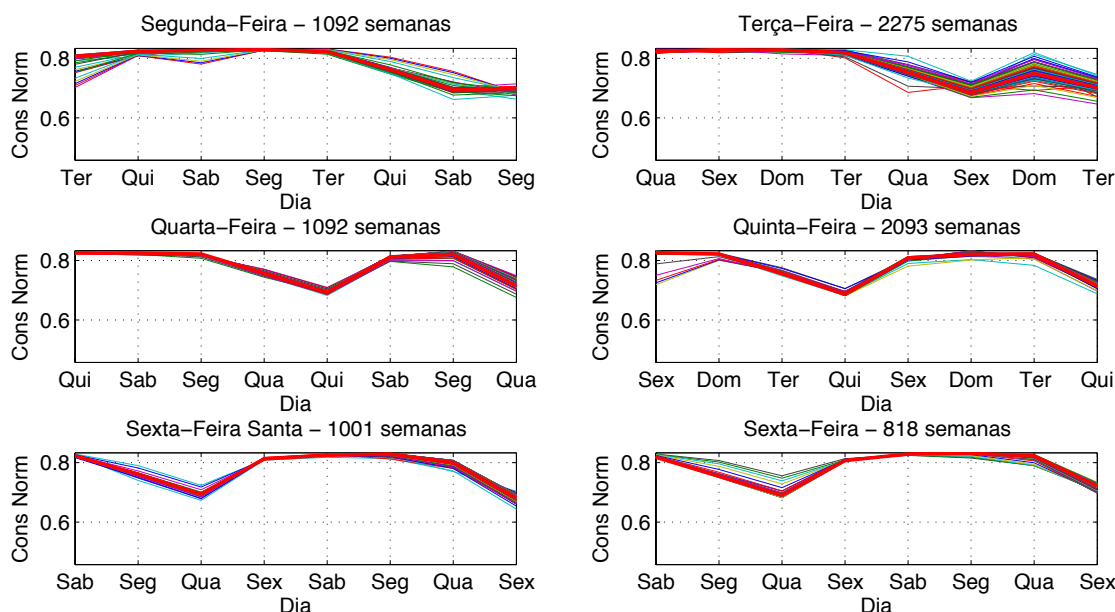
Ao contrário do que aconteceu com o ITL MS, a rede de *Kohonen* não conseguiu distinguir os padrões de consumo das sextas-feiras, como mostrado na Figura 5.5. Essa distinção entre os padrões de consumo dos feriados nas sextas-feiras representa um conhecimento adicional dado pelo algoritmo ITL MS.

Assim como o ITL MS, a rede de *Kohonen* também gerou *clusters* com casos anormais de consumo. Portanto, os resultados obtidos pela rede de *Kohonen* contribuem para validar os resultados gerados pelo ITL MS na clusterização do consumo nos feriados, sendo que com o ITL MS é possível gerar os valores virtuais de consumo, com a estratégia da Densificação, conforme visto no Capítulo 2.

#### 5.2.4 DENSIFICAÇÃO DOS DADOS

À medida que os dados vão sendo clusterizados, o algoritmo ITL MS gera os valores virtuais. A convergência do algoritmo foi alcançada depois de 54 iterações, onde 54 séries

virtuais foram geradas para cada amostra de dados reais (ver Figura 5.8). O número total de amostras virtuais em cada *cluster* é igual ao número de pontos reais multiplicadas pelo número de iterações realizadas pelo algoritmo.



**Figura 5.8** – Séries de dados virtuais geradas pelo ITL MS após 54 iterações, com 1092 (segunda-feira), 2275 (terça-feira), 1092 (quarta-feira), 2093 (quinta-feira), 1001 (sexta-feira santa) e 818 (sexta-feira) semanas virtuais cada. As linhas coloridas representam o consumo virtual e a linha vermelha mais grossa representa a moda de cada *cluster*.

### 5.2.5 PREDIÇÃO DE CONSUMO

A etapa de predição global da metodologia proposta neste trabalho é resultado da combinação de 2 modelos: (1) um algoritmo aplicado para ajustar o consumo em dias especiais (descrito na Seção 4.3.5); e (2) modelo responsável por prever padrões para consumo em dias normais (descrito na Seção 4.3.6). Em ambos os casos, apesar de diversos algoritmos de predição poderem ser usados, neste trabalho optou-se por usar RNA e RLM, por serem bastante empregados nesse domínio de aplicação, conforme visto no Capítulo 3.

#### 5.2.5.1 TRATAMENTO PARA OS FERIADOS

Nesta etapa, 6 redes neurais foram implementadas para modelar o comportamento de cada classe de ocorrência de feriados – uma para cada *cluster* identificado pelo ITL MS. O conjunto de entradas para cada rede neural corresponde aos valores de consumo dos 7 dias

que antecedem o feriado cujo consumo deseja ser “predito”, organizados conforme as entradas usadas para a clusterização com o ITL MS na Seção 5.2.3.

É importante ressaltar que o termo “predição” usado aqui se refere ao aprendizado do padrão de consumo dos feriados para que esse conhecimento seja aplicado, de modo a ajustar o modelo de predição global, conforme será visto na Seção 5.2.5.3.

Nesta metodologia, em virtude do pequeno número de amostras reais dos feriados disponíveis na base de dados da CELPA (considerados insuficiente para um adequado processo de treinamento e teste da RNA), um novo modelo foi adotado para realizar os estudos com os feriados, onde os dados virtuais foram utilizados no treinamento e os dados reais (de 2002 a 2012) reservados para serem usados como conjunto de teste.

A exemplo disso, no caso específico da CELPA, o maior *cluster* dos feriados possuía 25 amostras de séries de consumo antes da densificação. Se uma simples RNA com 8 entradas e 6 neurônios na camada escondida fosse aplicada, o número de pesos que deveriam ser ajustados estariam em torno de 48 (no melhor cenário). Dessa forma, haveria uma ausência de redundância de padrões para um treinamento adequado da rede e claramente haveria um número insuficiente de casos para validar e testar a capacidade de generalização da rede. Contudo, o uso dos valores virtuais tornou possível superar a insuficiência dos dados neste estudo de caso.

As melhores RNAs encontradas para cada *cluster* encontram-se caracterizadas na Tabela 5.1 (a última coluna representa a função de transferência para as camadas intermediárias). Cada RNA foi treinada usando 100 épocas.



Tabela 5.1 - Melhores RNAs

<i>Clusters</i>	Nº de séries para treino	Função de Treinamento	Nº de neurônios/camada	Função de Transferência
segunda-Feira	1092	Levenberg-Marquardt	8-6-1	Log-sigmoid/Linear
terça-Feira	2275	Levenberg-Marquardt	8-7-3-1	Hyperbolic tangent sigmoid/Linear/Linear
quarta-Feira	1092	Levenberg-Marquardt	8-7-3-1	Hyperbolic tangent sigmoid/Linear/Linear
quinta-Feira	2093	Levenberg-Marquardt	8-6-1	Linear/Linear
sexta-Feira Santa	1001	Levenberg-Marquardt	8-7-3-1	Hyperbolic tangent sigmoid/Log-sigmoid
sexta-Feira	818	BFGS quasi-Newton backpropagation	8-6-1	Hyperbolic tangent sigmoid/Linear

A avaliação do desempenho do algoritmo é baseada no Erro Percentual Médio Absoluto (MAPE), representado pela Equação 5.1, em relação aos valores de consumo predito comparados com os valores históricos. A Tabela 5.2 resume os resultados significantes obtidos.

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{X_i - \bar{X}_i}{X_i} \right| (100)}{n} \quad (5.1)$$

onde  $X_i$  é o valor de consumo real,  $\bar{X}_i$  é o valor do consumo previsto e  $n$  é o número de amostras de consumo.

Tabela 5.2 – Resultado da predição com RNA

Localização do feriado	Nº de séries para o teste ( $n$ )	MAPE	Desvio Padrão ( $s$ )	Intervalo de Confiança ( $CI$ )
segunda-Feira	12	0,83%	0,72%	(0,37%, 1,29%)
terça-Feira	25	1,74%	2,32%	(0,78%, 2,70%)
quarta-Feira	12	1,74%	1,36%	(0,87%, 2,61%)
quinta-Feira	23	1,10%	0,85%	(0,73%, 1,47%)
sexta-Feira	22	2,59%	1,94%	(1,68%, 3,50%)
TOTAL	94	1,65%	1,74%	(1,29%, 2,01%)

Os valores de erros encontram-se divididos em 5 classes, que estão ligados ao dia da semana em que o feriado ocorreu (em virtude dos *clusters* das sextas-feiras santas e sextas-feiras representarem o mesmo dia da semana, o erro MAPE para predizer uma sexta-feira é resultado da média dos resultados de suas RNAs). Os erros MAPE variaram de 0,83% (segunda-feira) a 2,59% (sexta-feira).

Além disso, uma análise estatística sobre o MAPE obtido foi conduzido para inferir um intervalo de confiança para o nível  $\alpha=0.05$ . Em virtude do número de amostras de teste

ser pequeno e a variância ser desconhecida, a função de distribuição de probabilidade t-Student foi utilizada para determinar o intervalo de confiança de  $CI = MAPE \pm t_{\alpha/2; q-1} s / \sqrt{q}$ , onde  $q$  é o número de amostras,  $s$  é o desvio padrão do erro e  $t_{\alpha/2; q}$  é o valor de  $t$  dentro de uma probabilidade acumulativa de  $1-\alpha$ , quando número de graus de liberdade foram  $q-1$ .

Os resultados mostram que no pior cenário, o erro máximo dentro do intervalo de confiança é de 3,5% para uma terça-feira. Geralmente, considerando todos os erros como uma única série (TOTAL), o erro MAPE foi de 1,65% com um valor máximo dentro do intervalo de confiança de 2,01%. Através da observação desses resultados, nota-se que, em geral, os modelos são mais precisos, quando comparados com os obtidos previamente obtidos com os modelos utilizados pela CELPA, que era em torno de 4%.

Para validar esse modelo, foram implementados uma RNA e uma RLM para prever dias especiais sem a utilização dos valores virtuais, ou seja, usando somente dados reais. A RNA e a RLM utilizaram o mesmo conjunto de dados (para treino e teste). Destes, a RNA obteve melhores resultados e suas características encontram-se descritas na Tabela 5.3.

Tabela 5.3 – RNA para predição de consumo sem densificação.

<b>Nº de Série de entradas para treino</b>	85
<b>Nº de Série de entradas para teste</b>	12
<b>Função de treinamento</b>	Levenberg-Marquardt
<b>Nº de neurônios por camada</b>	8-20-10-1
<b>Função de transferência</b>	Hyperbolic tangent sigmoid and Linear
<b>Nº de épocas</b>	100
<b>MAPE</b>	9,01%

A Tabela 5.3 mostra que a metodologia proposta (utilizando RNA e valores virtuais de consumo nos feriados) gerou um erro médio bem menor (1,65%) que o encontrado pelo modelo sem a densificação (9,01%). Isso prova a boa performance do ITL MS para aprender o padrão de consumo dos feriados e para gerar valores virtuais compatíveis com os valores reais.

### 5.2.5.2 PREDIÇÃO PARA OS DIAS NORMAIS

Uma rede neural foi implementada para prever o consumo dos dias normais e para representar a variação temporal (crescimento mensal/anual) do padrão de consumo de energia que foi perdida após as etapas de clusterização e densificação. Para efeito de comparação, uma RLM também foi implementada, com os mesmos conjuntos de entrada (não normalizados) para o treino e teste. O conjunto de treino foi o consumo dos anos de 2002 a 2011 e o de teste com os valores de consumo de 2012. As séries de entrada dos algoritmos foram organizadas da mesma forma que o modelo para os feriados.

Mais uma vez, a RNA obteve os melhores resultados com uma média de erro de 1,37%, desvio padrão de 1,2% e intervalo de confiança de (1,23%, 1,52%). O erro obtido com a RLM foi de 6,03%. A Tabela 5.4 mostra os resultados dos erros MAPE encontrados nessa análise.

Tabela 5.4 – Resumo da predição para dias normais

<b>Modelo</b>	<b>MAPE</b>
RNA	1,37 %
RLM	6,03%

Na Tabela 5.5, encontram-se descritos os parâmetros da melhor RNA implementada.

Tabela 5.5 – Melhor RNA

<b>Nº de séries para treino</b>	<b>Nº de séries para teste</b>	<b>Nº de épocas</b>	<b>Funções de Treinamento</b>	<b>Nº de neurônios/camada</b>	<b>Função de transferência</b>
3275	357	100	Levenberg Marquardt	8-6-1	Logsigmoid/Linear

### 5.2.5.3 PREDIÇÃO GLOBAL

O processo de predição global resulta da combinação dos modelos gerados para os dias normais e para os dias especiais. O objetivo desta abordagem é controlar o deslocamento da série e a sua tendência de crescimento, uma vez que o conhecimento obtido a partir das estratégias de clusterização e de densificação não possuem informação temporal sobre os feriados.

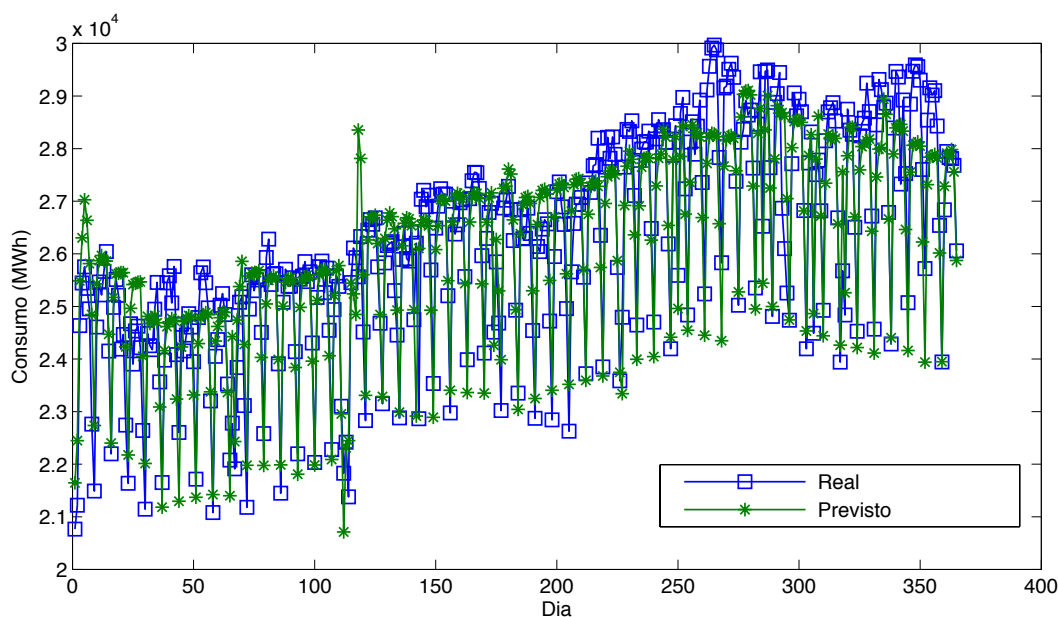
O teste consistiu na realização de previsões diárias para todo o ano de 2012 (tanto dias normais quanto feriados) usando somente valores históricos como entrada para as RNAs. O erro médio diário encontrado foi de 1,44% com desvio padrão de 1,36% e intervalo de confiança de (1,30%, 1,58%). A Figura 5.9 mostra os valores preditos e os valores reais obtidos nessa análise.

Com o objetivo de avaliar a eficiência estatística da metodologia proposta e também comparar seus resultados com os resultados dos modelos empregados pela concessionária, testes de confiança foram empregados. Para os testes de significância estatística, foi definida a hipótese de que o erro médio obtido pela metodologia proposta deveria ser menor que o erro obtido pelos modelos utilizados pela concessionária (em torno de 4%). Se a hipótese for nula ( $H_0: MAPE < 4\%$ ) então não teve evidência estatística suficiente para ser aceita, portanto ela é rejeitada e assume-se a hipótese alternativa ( $H_1: MAPE > 4\%$ ) é aceita, mostrando que a concessionária deve manter seus próprios métodos.

Dado um alto número de amostras ( $q=356$ ), uma distribuição normal foi usada para determinar a métrica de teste, sendo  $Z_{test} = \frac{MAPE-4\%}{s/\sqrt{q}}$ . Pela aplicação do nível de significância  $\alpha=0.05$ , o resultado  $Z_{test}$  foi comparado com o valor  $Z_\alpha$  obtido a partir da probabilidade acumulada de  $1-\alpha$ . Uma vez que  $Z_{test} = -35.28 < Z_\alpha = 1.65$ , evidência estatística suficiente é obtida para admitir que  $H_0$  deve ser aceita, isto é, com a probabilidade de 0.95, o uso da metodologia proposta será mais eficiente do que os modelos empregados pela concessionária.

Contudo, a eficiência do uso da metodologia fica mais aparente quando se analisa o modelo da previsão do consumo nos feriados treinados após a densificação (Tabela 5.2), que obteve um desempenho muito melhor quando comparado com os testes sem a densificação.

Para efeito de comparação, com a RLM, o erro MAPE foi de 6.28%. Similarmente, a hipótese de que o erro MAPE da metodologia proposta é muito menor do que o erro MAPE encontrado pela RLM foi testada. Nesse caso, observou-se que não há evidência estatística para rejeitar a hipótese. Isso deve-se ao fato de que para a RLM  $Z_{test} = \frac{MAPE-6.28\%}{s/\sqrt{q}} = -67.14$  é muito menor que  $Z_\alpha = 1.65$ .



**Figura 5.9** – Predição global de consumo de curto prazo

#### 5.2.5.4 TESTES PARA OUTROS HORIZONTES DE PREDIÇÕES

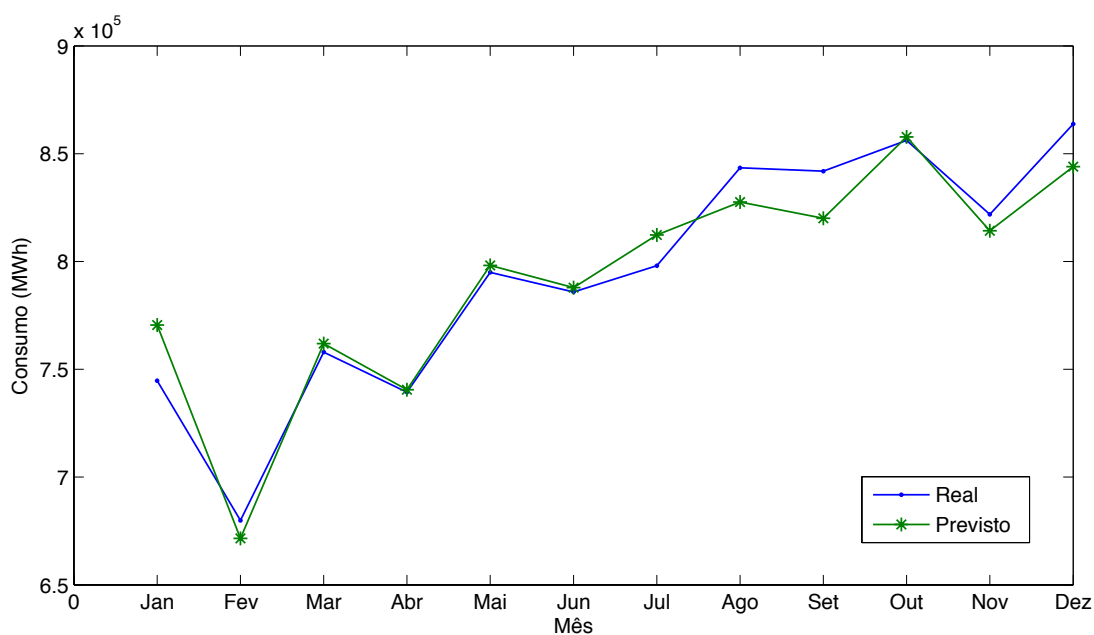
A metodologia foi aplicada em outros horizontes com o objetivo de avaliar o conhecimento adquirido sobre os dias especiais e determinar a aplicabilidade desta metodologia para ajustes em previsões de médio e longo prazo (DILHANI; JEENANUNTA, 2016). A utilização da metodologia para outros horizontes de predição é resultado de sucessivas previsões diárias conforme o período a ser previsto. Para previsões mensais, por exemplo, a metodologia seria executada para 30 dias à frente e seus erros diários seriam acumulados para se chegar ao erro final da predição.

Em virtude do desempenho da predição ser o resultado da soma acumulada de sucessivos erros de previsões diárias, uma de longo prazo (predição anual) não é considerada viável para essa metodologia pois seus erros cumulativos gerariam valores muito elevados.

Para a análise de médio prazo (predição mensal), todos os meses do ano de 2011 foram preditos através da utilização dos valores de consumo diário. Os resultados obtidos com essa análise, cujos erros encontram-se demonstrados na Tabela 5.6, são considerados satisfatórios, onde a média do erro foi 1,3% (também abaixo do valor encontrado pela concessionária). A Figura 5.10 mostra os valores preditos e os valores reais utilizados nessa análise.

Tabela 5.6 – Resultados da predição mensal

Mês	MAPE
janeiro	3,46%
fevereiro	1,22%
março	0,52%
abril	0,14%
maio	0,41
junho	0,25%
julho	1,78%
agosto	1,88%
setembro	2,59%
outubro	0,19%
novembro	0,92%
dezembro	2,29%



**Figura 5.10** – Consumo real x consumo previsto para o horizonte de médio prazo

## 5.2.6 DESEMPENHO E ESCALABILIDADE

Além de validar a qualidade dos resultados da predição obtidos pela metodologia proposta, é importante também avaliarmos o desempenho computacional de seus algoritmos,

para que seus resultados sejam gerados em um tempo adequado que possibilitem tomadas de decisão pela concessionária.

Na prática, a eficiência dos algoritmos pode ser medida através do cálculo do tempo médio das seguintes etapas da metodologia: clusterização, densificação e predição global (incluindo as etapas de ajuste dos dias especiais e predição para os dias normais, e suas respectivas fases de treino e teste). A Tabela 5.7 mostra os tempos de execução, em segundos, dessas etapas.

Tabela 5.7 – Tempos de Execução

<b>Etapas</b>	<b>Tempo (s)</b>
Clusterização	12,37
Densificação	10,03
Predição Global	110,67
TOTAL	133,07

Por meio da análise dos tempos de execução, é possível verificar que a metodologia é capaz de gerar resultados em um tempo considerado satisfatório para tomadas de decisão quando se trata de uma análise de predição para o próximo dia. Adicionalmente, uma outra característica importante é a avaliação da escalabilidade do algoritmo. Nesse caso, duas situações devem ser analisadas: (1) a aplicação da metodologia em outra base diária de consumo; e (2) o uso da metodologia para prever o consumo em outros horizontes. No primeiro caso, em virtude de não haver uma outra base de dados para teste, a metodologia não pôde ser testada, mas em uma análise preliminar, sua aplicação pode ser perfeitamente adaptada à outras bases de dado de consumo diário. No segundo caso, conforme visto na Seção 5.2.5.3, a metodologia alcançou bons resultados para o horizonte de médio prazo. Contudo, sua aplicação no horizonte de longo prazo é considerado inadequado.

### **5.3 CONSIDERAÇÕES FINAIS**

Neste capítulo, a metodologia proposta foi aplicada ao estudo de caso da CELPA, onde cada uma de suas etapas foram detalhadas e os seus resultados foram apresentados.

A partir da análise dos resultados apresentados, constatou-se que o emprego dessa metodologia no cenário de predição de energia é considerada satisfatória.

No capítulo a seguir, serão apresentadas as conclusões deste trabalho, bem como as principais contribuições dessa metodologia.



## 6. CONCLUSÕES

### 6.1 CONSIDERAÇÕES DA TESE

Neste trabalho, foi proposta uma nova metodologia para predição de curto prazo, cujo foco é a sua aplicação em bases de dados com amostras relacionadas à ocorrência de eventos pontuais, como acontece com o consumo em dias especiais. Com a constatação de que os eventos pontuais de consumo possuem informações importantes e que não podem ser desprezadas, a metodologia desenvolvida neste trabalho se propõem a avaliar o efeito que esses dias (especificamente os feriados) tem sobre a série de dados de consumo.

A metodologia trata os dias especiais através da aplicação do algoritmo ITL MS para aprender o padrão de consumo desses dias (por meio da clusterização do consumo), bem como estender a quantidade de amostras de consumo nos feriados presentes na base dados (geradas através do processo de Densificação). O conhecimento adquirido através do ITL MS é posteriormente utilizado em um modelo de predição para validação da eficiência da metodologia.

Diante disso, a metodologia apresentada neste trabalho tratou dois problemas específicos: (a) a influência que o consumo sofre em relação ao dia da semana em que o feriado ocorre; e (b) bases de dados com variáveis escassas. A combinação dos processos de clusterização e densificação geradas pelo algoritmo ITL MS estabelece uma solução para esses problemas, cujos resultados serviram de entrada para uma RNA que foi treinada com os valores virtuais (gerados pela densificação) e testada com os dados reais da base de dados, objetivando melhorar a eficiência da predição.

Ressalta-se que na literatura, apesar de ser possível encontrar alguns trabalhos que consideram os feriados em estratégias de predição de energia, o conhecimento sobre o feriados não é abordado da mesma forma que o proposto aqui com os processos de clusterização e densificação.

### 6.2 CONTRIBUIÇÕES

A partir dos resultados obtidos com a aplicação da metodologia, as seguintes contribuições podem ser elencadas:

- O ITL MS provou ser uma eficiente e não empírica estratégia para descoberta de padrões em cenários elétricos. Esse conhecimento deveu-se especificamente à

aplicação dos processos de clusterização, que identifica os padrões existentes na base dados e ao processo de densificação, que gera valores virtuais para bases de dados com valores escassos. Os resultados do ITL MS podem ser empregados em diversas estratégias de mineração de dados, como os modelos de predição de consumo;

- A densificação, que gerou valores virtuais ao longo da aplicação da metodologia, provou ser um dos diferenciais desta tese, pois através dela foi possível aumentar a capacidade do modelo de predição aprender os relacionamentos existentes entre os valores de consumo da série estudada;
- A metodologia alcançou bons resultados com a combinação dos modelos de predição para dias especiais e para dias normais, onde os erros de predição foram inferiores ao erros obtidos pelos modelos utilizados pela concessionária de energia estudada, que eram em torno de 4%. Para qualquer concessionária, erros menores de predição possibilitam substancial economia financeira em suas estratégias de planejamento do sistema elétrico;
- A metodologia desenvolvida também pode ser estendida para realizar predições de médio prazo, com bons resultados alcançados nos estudos desenvolvidos neste trabalho;
- A possibilidade de aplicação da metodologia para descoberta e tratamento de outros fenômenos inerentes ao setor elétrico, como a ocorrência de apagões;
- Por se tratar de uma metodologia com alto grau de generalização, sua utilização pode contribuir com estudos em outros cenários do setor elétrico, tais como operação, manutenção, supervisão e novos serviços em ambientes *smart grids*, onde a existência de variáveis com poucas amostras pode ser um gargalo no desenvolvimento de seus modelos.
- Geração de subsídios ao desenvolvimento de projetos de pesquisa, desenvolvimento e inovação, junto às concessionárias, na área de pesquisa em questão;
- Consolidação da cooperação com outros grupos de pesquisa no país e no exterior, tais como o INESC -TEC, Portugal, e o INPE, Brasil, a partir das atividades envolvidas nesta proposta e desenvolvidas conjuntamente com os pesquisadores daqueles Institutos.
- Divulgação das ideias e dos resultados obtidos, fruto do desenvolvimento desta tese, à comunidade, a partir da publicação do artigo:

- REGO, L. P.; SUMAILI, J.; MIRANDA, V.; FRANCES, Carlos Renato Lisboa; SILVA, Marcelino Silva da; SANTANA, A. L.; MEAN SHIFT DENSIFICATION OF SCARCE DATA SETS IN SHORT- TERM ELECTRIC POWER LOAD FORECASTING FOR SPECIAL DAYS. ELECTRICAL ENGINEERING, 2016.

### **6.3 TRABALHOS FUTUROS E POSSÍVEIS DESDOBRAMENTOS**

A seguir, são apresentadas algumas sugestões para trabalhos futuros que podem ser desenvolvidos através da metodologia proposta nesta tese:

- Implementar outras técnicas de densificação e clusterização, com vistas a possibilitar comparações com as técnicas adotadas neste trabalho;
- Analogamente, utilizar outros métodos de predição, no sentido de aferir a acurácia dos métodos atualmente adotados;
- Conceber e elaborar interface amigável para facilitar e estimular a utilização dos algoritmos implementados, no âmbito das concessionárias de energia;
- Aplicar a metodologia em outros cenários do setor elétrico, mais especificamente na área de *smart grids*, em estudos de previsão relacionados com o consumo pelo lado da demanda;
- Aplicar o ITL MS para estudo de outras variáveis que podem influenciar sobremaneira nas análises de consumo *spot*, como por exemplo o estudo do perfil de consumo por tipo de consumidor e sua utilização em uma estratégia de previsão de consumo por classe;
- Estabelecer metas concretas no sentido de patentear o arcabouço de conhecimentos e algoritmos desenvolvidos no âmbito desta tese.

### **6.4 DIFICULDADES ENCONTRADAS**

Ao longo do desenvolvimento deste trabalho, foram encontradas diversas dificuldades para a sua execução. Relatá-las aqui, podem servir de subsídio para pesquisadores e estudantes que tenham interesse em utilizar os conhecimentos disponíveis neste trabalho.

As principais dificuldades encontradas, foram:

- Necessidade de um exaustivo estudo para apropriação dos modelos de inteligência computacional e do ITL MS, que possuem um forte teor matemático;

- Dificuldades com a implementação das técnicas utilizadas, bem como dos algoritmos necessários para formatação dos dados e dos algoritmos que integram os modelos existentes na metodologia;
- Intensivo estudo da literatura e de trabalhos dos grupos de pesquisa para concepção da proposta da tese, bem como para identificação de suas contribuições;
- Obtenção da base de dados de consumo real para realização dos estudos da metodologia implementada;
- Obtenção de apoio dos especialistas do domínio do setor elétrico para auxiliar na análise dos resultados obtidos;

## BIBLIOGRAFIA

- AGARWAL, Paridhi; SAO, Anil; BHAVSAR, Arnav. Mean-shift based segmentation of cell nuclei in cervical PAP-smear images. In: **Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2015 Fifth National Conference on**. IEEE, 2015. p. 1-4.
- ANDERSEN, Thomas J.; WILAMOWSKI, Bogdan M. A. Modified Regression Algorithm for Fast One Layer Neural Network Training. In: **World Congress of Neural Networks**. 1995. p. 687-690.
- ARORA, Siddharth; TAYLOR, James W. Short-term forecasting of anomalous load using rule-based triple seasonal methods. **IEEE transactions on Power Systems**, v. 28, n. 3, p. 3235-3242, 2013.
- BATTITI, Roberto. First-and second-order methods for learning: between steepest descent and Newton's method. **Neural computation**, v. 4, n. 2, p. 141-166, 1992.
- COLE, Rowena Marie. **Clustering with genetic algorithms**. University of Western Australia, 1998.
- COMANICIU, Dorin; MEER, Peter. Mean shift analysis and applications. In: **Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on**. IEEE, 1999. p. 1197-1203.
- CHO, Haeran et al. Modelling and forecasting daily electricity load via curve linear regression. In: **Modeling and Stochastic Learning for Forecasting in High Dimensions**. Springer International Publishing, 2015. p. 35-54.
- CHO, Hanjoo; KIM, Young Hwan. Image segmentation using linked mean-shift vectors with region attribution analysis. In: **Ph. D. Research in Microelectronics and Electronics (PRIME), 2015 11th Conference on**. IEEE, 2015. p. 188-191.
- CHUNSHAN, Yang; XIAOFENG, Li. Study and application of data mining and NARX neural networks in load forecasting. In: **Computer Science and Network Technology (ICCSNT), 2015 4th International Conference on**. IEEE, 2015. p. 360-364.
- CUGLIARI, Jairo; GOUDE, Yannig; POGGI, Jean-Michel. Disaggregated electricity forecasting using wavelet-based clustering of individual consumers. In: **Energy Conference (ENERGYCON), 2016 IEEE International**. IEEE, 2016. p. 1-6.
- CYBENKO, G. Continuous valued neural networks with two hidden layers are sufficient, Department of computer science. **Trfts. University**, 1988.

- DILHANI, MHMR Shyamali; JEENANUNTA, Chawalit. Daily electric load forecasting: Case of Thailand. In: **2016 7th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES)**. IEEE, 2016. p. 25-29.
- FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. From data mining to knowledge discovery in databases. **AI magazine**, v. 17, n. 3, p. 37, 1996.
- FIDALGO, J. N.; LOPES, JA Pecas. Load forecasting performance enhancement when facing anomalous events. **IEEE transactions on power systems**, v. 20, n. 1, p. 408-415, 2005.
- FREDERICO DIAS DINIZ, Carlos. Sistema de previsão de consumo de energia elétrica em curto prazo em função da variável temperatura utilizando redes neurais. 2011.
- FUKUNAGA, Keinosuke; HOSTETLER, Larry. The estimation of the gradient of a density function, with applications in pattern recognition. **IEEE Transactions on information theory**, v. 21, n. 1, p. 32-40, 1975.
- GALL, Norman. Apagão na política energética. **Braudel Papers**, v. 32, 2002.
- GARULLI, Andrea; PAOLETTI, Simone; VICINO, Antonio. Models and techniques for electric load forecasting in the presence of demand response. **IEEE Transactions on Control Systems Technology**, v. 23, n. 3, p. 1087-1097, 2015.
- GUERINI, Alice; DE NICOLAO, Giuseppe. Long-and short-term electric load forecasting on quarter-hour data: A 3-torus approach. In: **Environment and Electrical Engineering (EEEIC), 2016 IEEE 16th International Conference on**. IEEE, 2016. p. 1-4.
- HAGAN, Martin T.; MENHAJ, Mohammad B. Training feedforward networks with the Marquardt algorithm. **IEEE transactions on Neural Networks**, v. 5, n. 6, p. 989-993, 1994.
- HAN, Jiawei; PEI, Jian; KAMBER, Micheline. **Data mining: concepts and techniques**. Elsevier, 2011.
- HAND, David J.; MANNILA, Heikki; SMYTH, Padhraic. **Principles of data mining**. MIT press, 2001.
- HASHMI, Md Umar; ARORA, Varun; PRIOLKAR, Jayesh G. Hourly electric load forecasting using Nonlinear AutoRegressive with eXogenous (NARX) based neural network for the state of Goa, India. In: **Industrial Instrumentation and Control (ICIC), 2015 International Conference on**. IEEE, 2015. p. 1418-1423.
- HAYKIN, Simon; NETWORK, Neural. A comprehensive foundation. **Neural Networks**, v. 2, n. 2004, 2004.

- HONG, Tao; FAN, Shu. Probabilistic electric load forecasting: A tutorial review. **International Journal of Forecasting**, v. 32, n. 3, p. 914-938, 2016.
- JAIN, Anil K.; DUBES, Richard C. **Algorithms for clustering data**. Prentice-Hall, Inc., 1988.
- KASABOV, Nikola K. **Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering**, A Bradford Book. 1998.
- KILIÇ, Volkan et al. Mean-Shift and Sparse Sampling Based SMC-PHD Filtering for Audio Informed Visual Speaker Tracking. **IEEE Transactions on Multimedia**, v. PP, p. 1-1, 2016.
- KOHONEN, Teuvo. The 'neural' phonetic typewriter. **Computer**, v. 21, n. 3, p. 11-22, 1988.
- LAHOUAR, Ali; SLAMA, Jaleddine Ben Hadj. Random forests model for one day ahead load forecasting. In: **Renewable Energy Congress (IREC), 2015 6th International**. IEEE, 2015. p. 1-6.
- LEMOS, Alan Alexander Mendes; MYNBAEV, Kairat Turysbekovich. **Manual de Econometria**. FGV Editora, 2004.
- LINO, P. et al. Energy tariff and demand response in Brazil: An analysis of recent proposals from the regulator. In: **Innovative Smart Grid Technologies (ISGT Latin America), 2011 IEEE PES Conference on**. IEEE, 2011. p. 1-5.
- LUZ, Cristina Pimenta de Mello Spinetti; GOMES, Leonardo Lima; BRANDÃO, Luiz Eduardo Teixeira. Análise da Dinâmica do Mercado a Termo de Energia Elétrica no Brasil. **Revista Brasileira de Gestão de Negócios**, v. 14, n. 44, p. 314-334, 2012.
- LUTHULI, Qiniso W.; FOLLY, Komla A. Short term load forecasting using artificial intelligence. In: **PowerAfrica, 2016 IEEE PES**. IEEE, 2016. p. 129-133.
- MANNILA, Heikki. Data Mining: Machine Learning, Statistics, and Databases. In: **ssdbm**. 1996. p. 2-9.
- MIRANDA, Vladimiro; CASTRO, Adriana R. Garcez; LIMA, Shigeaki. Diagnosing faults in power transformers with autoassociative neural networks and mean shift. **IEEE Transactions on Power Delivery**, v. 27, n. 3, p. 1350-1357, 2012.
- MOTLAGH, Omid; GROZEV, George; PAPAGEORGIOU, Elpiniki I. A Neural Approach to Electricity Demand Forecasting. In: **Artificial Neural Network Modelling**. Springer International Publishing, 2016. p. 281-306.
- MULYADI, Y. et al. Anomalous STLF for Indonesia power system using Artificial Neural Network. In: **Science and Technology (TICST), 2015 International Conference on**. IEEE, 2015. p. 1-4.

- MURE, Simon; GRENIER, Thomas; BENOIT-CATTIN, Hugues. Unsupervised spatiotemporal video clustering a versatile mean-shift formulation robust to total object occlusions. In: **2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. IEEE, 2016. p. 1536-1540.
- PANAPAKIDIS, I. P.; PAPAGIANNIS, G. K. Application of the load profiling methodology in short-term bus load forecasting. In: **MedPower 2014**. IET, 2014. p. 1-8.
- PANAPAKIDIS, Ioannis P.; CHRISTOFORIDIS, George C.; PAPAGIANNIS, Grigoris K. Hybrid computational intelligence model for Short-Term bus load forecasting. In: **Environment and Electrical Engineering (EEEIC), 2015 IEEE 15th International Conference on**. IEEE, 2015. p. 2029-2034.
- PAPADOPOULOS, Sokratis; KARAKATSANIS, Ioannis. Short-term electricity load forecasting using time series and ensemble learning methods. In: **Power and Energy Conference at Illinois (PECI), 2015 IEEE**. IEEE, 2015. p. 1-6.
- PAPARODITIS, Efstathios; SAPATINAS, Theofanis. Short-term load forecasting: the similar shape functional time-series predictor. **IEEE Transactions on Power Systems**, v. 28, n. 4, p. 3818-3825, 2013.
- PARZEN, Emanuel. On estimation of a probability density function and mode. **The Annals of Mathematical Statistics**, v. 33, n. 3, p. 1065–1076, 1962.
- QUILUMBA, Franklin L. et al. Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities. **IEEE Transactions on Smart Grid**, v. 6, n. 2, p. 911-918, 2015.
- RAO, Sudhir et al. Information theoretic mean shift algorithm. In: **2006 16th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing**. IEEE, 2006. p. 155-160.
- RAZA, Muhammad Qamar et al. A comparative analysis of PSO and LM based NN short term load forecast with exogenous variables for smart power generation. In: **Intelligent and Advanced Systems (ICIAS), 2014 5th International Conference on**. IEEE, 2014. p. 1-6.
- RÉNYI, Alfréd. Some fundamental questions of information theory. **Selected Papers of Alfred Renyi**, v. 2, n. 174, p. 526-552, 1976.
- REZENDE, Solange Oliveira. **Sistemas inteligentes: fundamentos e aplicações**. Editora Manole Ltda, 2003.
- SAIPUNYA, Sarunrut; THEERA-UMPON, Nipon; AUEPHANWIRIYAKUL, Sansanee.



Cross-substation short-term load forecasting based on types of customer usage characteristics. In: **Information and Communication Technology, Electronic and Electrical Engineering (JICTEE), 2014 4th Joint International Conference on.** IEEE, 2014. p. 1-5.

SANTANA, Á. L. de, Estratégias para a Melhoria da modelagem e interpretabilidade de redes bayesianas; Tese (Doutorado), Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, Belém, 2008.

SHAO, Zhen et al. Multivariate statistical and similarity measure based semiparametric modeling of the probability distribution: A novel approach to the case study of mid-long term electricity consumption forecasting in China. **Applied Energy**, v. 156, p. 502-518, 2015.

SINDHUJA, Gandham. Comparative analysis of mean shift in object tracking. In: **Power, Control, Communication and Computational Technologies for Sustainable Growth (PCCCTSG), 2015 Conference on.** IEEE, 2015. p. 283-287.

SULAIMAN, S. M.; JEYANTHY, P. Aruna; DEVARAJ, D. Artificial neural network based day ahead load forecasting using Smart Meter data. In: **Power and Energy Systems: Towards Sustainable Energy (PESTSE), 2016 Biennial International Conference on.** IEEE, 2016. p. 1-6.

SUMAILI, Jean et al. On the use of information theoretic mean shift for electricity load patterns clustering. In: **PowerTech, 2011 IEEE Trondheim.** IEEE, 2011. p. 1-6.

SYMEONIDIS, A. L.; MITKAS, P. A. Agent Intelligence through Data Mining. Multi-agent Systems, Artificial Societies, and Simulated Organizations Series 14: 200. USA: International Book Series. 2005.

TENFEN, D. et al. Microgrids and Microgeneration in Brazilian Energy Market: a Discussion of Regulatory and Comercial Aspects. In: **9º Latin-American Congress on Electricity Generation and Transmission (CLAGTEE).** 2013.

VELASCO, Lemuel Clark P. et al. Next day electric load forecasting using Artificial Neural Networks. In: **Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), 2015 International Conference on.** IEEE, 2015. p. 1-6.

- WANG, Xin et al. Factors that impact the accuracy of clustering based load forecasting. In: **Industry Applications Society Annual Meeting, 2015 IEEE**. IEEE, 2015. p. 1-6.
- XU, Rui; WUNSCH, Donald. Survey of clustering algorithms. **IEEE Transactions on neural networks**, v. 16, n. 3, p. 645-678, 2005.
- ZHANG, Guoqiang; PATUWO, B. Eddy; HU, Michael Y. Forecasting with artificial neural networks:: The state of the art. **International journal of forecasting**, v. 14, n. 1, p. 35-62, 1998.
- WANG, Zhenyu; ZHENG, Guilin. The application of mean-shift cluster in residential appliance identification. In: **Control Conference (CCC), 2011 30th Chinese**. IEEE, 2011. p. 3111-3114.
- ZHOU, Jia-Xiang; LI, Zhi-Wei; FAN, Chong. Improved fast mean shift algorithm for remote sensing image segmentation. **IET Image Processing**, v. 9, n. 5, p. 389-394, 2014.